A probabilistic representation of the solution of some quasi-linear PDE with a divergence form operator. Application to existence of weak solutions of FBSDE

Antoine Lejay

A Probabilistic Representation of the Solution of some Quasi-Linear PDE with a Divergence Form Operator. Application to Existence of Weak Solutions of FBSDE

Antoine Lejay[1] — Mathematical Institute (Oxford University)

Abstract: We extend some results on time-homogeneous processes generated by divergence form operators to time-inhomogeneous ones. These results concern the decomposition of such processes as Dirichlet process, with an explicit expression for the term of zero-quadratic variation. Moreover, we extend some results on the Itô formula and BSDEs related to weak solutions of PDEs, and we study the case of quasi-linear PDEs. Finally, our results are used to prove the existence of weak solutions to Forward-Backward Stochastic Differential Equations (FBSDEs).

Keywords: quasi-linear PDE, divergence form-operator, Forward-Backward Stochastic Differential Equation, time reversal of a diffusion, Dirichlet process

AMS Classification: 60H10 (35J60 35R60 60J60)

Financial support: Partially supported by European Union's TMR Stochastic Analysis Network (grant reference 960075)

Published in *Stochastic Processes and their Applications.* **110**:1, pp. 145–176, 2004

Archives, links & reviews:

- DOI: 10.1016/j.spa.2003.09.012
- MR Number: 2052140

[1]Current address: Projet OMEGA (INRIA Lorraine)
IECN, Campus scientifique
BP 239
54506 Vandœuvre-lès-Nancy CEDEX (France)
E-mail: `Antoine.Lejay@iecn.u-nancy.fr`

 38 pages

Introduction

In this article we study the connection between non-linear PDEs and stochastic processes generated by divergence form operators. This link is done using the theory of Backward Stochastic Differential Equations (BSDEs). The differential operators we consider are of type

$$L = \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{i,j}(t,x)\frac{\partial}{\partial x_j}\right) + b_i(t,x)\frac{\partial}{\partial x_i} \tag{1}$$

where a is uniformly elliptic, and a and b are bounded. Using the property of the fundamental solution of $\frac{\partial}{\partial t}+L$, it is easily proved that L is the infinitesimal generator of a continuous, stochastic process $(X, \mathbb{P}_{s,y})$.

Before studying BSDEs and some applications, we show that X is a Dirichlet process in the sense of Föllmer (See Föllmer, 1981) under $\mathbb{P}_{s,y}$ for any starting point, that is $X_t = y + M_t + V_t$, where M is a local martingale and V_t a term of zero-quadratic variation. This means that for any $C > 0$,

$$\lim_{\substack{\mathrm{mesh}(\Pi)\to 0\\ \Pi=\{\,0\leq t_1\leq\dots\leq t_k\leq T\,\}}} \mathbb{P}_{s,y}\left[\sum_{i=1}^{k-1}(V_{t_{i+1}} - V_{t_i})^2 > C\right] = 0.$$

Although X is not in general a semi-martingale, this result allows to define the martingale part M of X, for which a martingale representation theorem holds. However, there are different possibilities to characterize the process V: See for example Ōshima, 1992a,b; Rozkosz, 2002. Our result uses an explicit decomposition of X as

$$\begin{aligned} X_t = y &+ \frac{1}{2}M_t + \frac{1}{2}(\overline{M}_{T-t} - \overline{M}_T) \\ &- \frac{1}{2}\int_s^t \Gamma^{-1}a\cdot\nabla\Gamma(s,y,r,X_r)\,\mathrm{d}r + \int_s^t b(r,X_r)\,\mathrm{d}r,\ t\in[s,T],\ \mathbb{P}_{s,y}, \end{aligned} \tag{2}$$

where $\overline{M}$ is a martingale with respect to the filtration generated by the "future" of X_t, that is $\sigma(X_r, r\in[t,T])$. These results were already known for processes generated by time-homogeneous divergence form operators: See Rozkosz and Slomiński, 1998; Lyons and Stoica, 1999, ... One of the main interest of the decomposition (2) is that it allows to define some stochastic integrals driven by X: See Rozkosz, 1996a; Rozkosz and Slomiński, 1998; Lyons and Stoica, 1999; Lejay, 2002b. As an application, we prove a linear Feynman-Kac formula for the semi-group of the differential operator $A = L^0 + b_i\partial_{x_i} + c - \partial_{x_i}(d_i\cdot)$, where $L^0 = \frac{1}{2}\partial_{x_i}(a_{i,j}\partial_{x_j})$. Let $(X, \mathbb{P}_{s,y}; (s,y) \in$

$[0,T]\times\mathbb{R}^N$) be the process generated by L^0. Let $(P_{s,t})_{t\geq s}$ be the semi-group generated by A, with transition density function Υ. Then for almost every x,

$$\begin{aligned} P_{s,T}g(x) = \int_{\mathbb{R}^N} \Upsilon(s,y,t,x)g(x)\,\mathrm{d}x = \mathbb{E}_{s,y}\Bigg[\exp\Bigg(\int_s^T a^{-1}b(r,X_r)\,\mathrm{d}M_r \\ + \int_s^T \overline{a}^{-1}\overline{d}(r,X_r)\,\mathrm{d}\overline{M}_r - \frac{1}{2}\int_s^T a^{-1}(b-d)\cdot(b-d)(r,X_r)\,\mathrm{d}r \\ + \int_s^T c(r,X_r)\,\mathrm{d}r + \int_s^T \Gamma^{-1}d\cdot\nabla\Gamma(s,y,r,X_r)\,\mathrm{d}r\Bigg)g(X_T)\Bigg], \end{aligned} \tag{3}$$

where Γ is the transition density function of L^0. This result extends the one of Lunt et al., 1998, where a formula was provided for $\int_{\mathbb{R}^N} P_{s,T}g(x)f(x)\,\mathrm{d}x$.

The solutions of semi-linear PDEs of type

$$\begin{cases} \frac{\partial u(t,x)}{\partial t} + Lu(t,x) + h(t,x,u(t,x),\nabla u(t,x)) = 0,\ (t,x)\in[0,T)\times\mathcal{O}, \\ u(t,x)=0 \text{ on } [0,T)\times\partial\mathcal{O} \text{ and } u(T,x)=g(x) \text{ on } \mathcal{O}, \end{cases} \tag{4}$$

are generally weak solution, *i.e.*, $u(t,\cdot)$ belongs to the Sobolev space $\mathrm{H}^1_0(\mathcal{O})$. But, it is still true that $(Y,Z) = (u(t,X_t),\nabla u(t,X_t))_{t\in[s,T]}$ is the solution of the BSDE, $\mathbb{P}_{s,y}$-almost surely, for any $t\in[s,t]$,

$$Y_t = g(X_T)\mathbf{1}_{\{T\leq\tau\}} + \int_{t\wedge\tau}^{T\wedge\tau} h(r,X_r,Y_r,Z_r)\,\mathrm{d}r - \int_{t\wedge\tau}^{T\wedge\tau} Z_r\,\mathrm{d}M_r \tag{5}$$

where τ is the first exit time of X from $\mathcal{O}$. As Y is adapted to the filtration generated by X, Y_s is deterministic and is equal to $u(s,y)$. Then, the theory of BSDEs may be applied to weak solutions of PDEs, and not only to classical and viscosity solutions, as it was proved first, under additional regularity assumptions, in Barles and Lesigne, 1997 and then in Lejay, 2002a; Bally et al., 2005; Stoica, 2003; Rozkosz, 2003. In this article, we extend the results of Lejay, 2002a to time-inhomogeneous processes, but we also give some precisions about the starting points (s,y) for which (5) holds, when h belongs only to $\mathrm{L}^{2,2}(0,T;\mathcal{O})$. This could also lead to a better understanding of the Itô formula for processes generated by divergence form operators. Besides, we prove that there exists a version $\widehat{u}$ of the solution of (4) such that $t\mapsto\widehat{u}(t,X_t)$ is continuous under $\mathbb{P}_{s,y}$ for almost every (s,y), although $\widehat{u}$ may fail to be continuous. For time-homogeneous operators and elliptic PDEs, this could follow from potential theory (quasi-continuity, ...): See Fukushima et al. (1994) for example. Although our result is more restrictive than the ones provided by potential theory, we do not need here to define some capacity.

Afterwards, we explain how this result could be used for quasi-linear PDEs, that is when the coefficients of L are themselves dependent on the solution:

$$L = L^u = \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{i,j}(t,x,u(t,x))\frac{\partial}{\partial x_j}\right) + b_i(t,x,u(t,x),\nabla u(t,x))\frac{\partial}{\partial x_i}.$$

In this case, it is almost immediate that a weak solution is also a mild solution, that is

$$u(s,y) = P^u_{s,T}g(y) + \int_s^T P^u_{s,r}h(r,x,u(r,x),\nabla u(r,x))\,\mathrm{d}r$$

when (s,y) belong to set of points that depends only on $h(\cdot,\cdot,0,0)$. We also prove that a mild solution is a weak solution.

If L is a quasi-linear differential non-divergence form operator, then the solution $u(s,y)$ of (4), if it is unique, may be found as Y_s, where (X,Y,Z) is the solution of the Forward-Backward Stochastic Differential Equation (FBSDE):

$$\begin{cases} X_t = y + \int_s^t \sigma(r,X_r,Y_r)\,\mathrm{d}B_r + \int_s^t b(r,X_r,Y_r,Z_r)\,\mathrm{d}r, \\ Y_t = g(X_T) + \int_t^T h(r,X_r,Y_r,Z_r)\,\mathrm{d}r - \int_t^T Z_r\,\mathrm{d}B_r, \end{cases} \tag{6}$$

for $t \in [s,T]$, $\mathbb{P}_{s,y}$-almost surely. Here B is a Brownian motion, $\sigma\sigma^{\mathrm{T}} = a$, and $L^u = \frac{1}{2}a_{i,j}(t,x,u(t,x))\frac{\partial^2}{\partial x_i \partial x_j} + b_i(t,x,u(t,x),\nabla u(t,x))\frac{\partial}{\partial x_i}$. It is hopeless to expect that such a representation holds for divergence form operators, since X is in general not a semi-martingale.

Although it is possible to consider FBSDEs without any reference to quasi-linear PDEs, using PDEs may be helpful. The book Ma and Yong, 1999 contains a review of results on FBSDEs. At the best of our knowledge, excepted in Antonelli and Ma, 2002, solutions of FBSDEs of type (6) have always been considered as strong solutions. This means that, as for SDEs, the Brownian motion B is given first, and then X, Y and Z are adapted to its natural filtration. This requires some strong assumptions on the coefficients. Mainly, $(t,x) \mapsto b(t,x,\cdot,\cdot)$ and $(t,x) \mapsto h(t,x,\cdot,\cdot)$ shall be Lipschitz continuous. A natural question is to know if there exists some non-trivial weak solutions for (6), where only the distribution of the process (B,X,Y,Z) is specified.

One may first think to apply the same methods as for proving the existence of weak solutions of SDEs. However, the martingale problem seems not to be easy to state. The Girsanov theorem adds a drift term on each part of the system. So, it could not be used in our case to add a drift $\int_s^t b(r,X_r,Y_r,Z_r)\,\mathrm{d}r$ to X, otherwise it means that a term like $\int_s^t Z_r b(r,X_r,Y_r,Z_r)\,\mathrm{d}r$

is already present in the expression of Y. Moreover, a direct proof using some approximations by strong solutions of FBSDE (See for example Rozkosz and Słomiński (1991) for SDEs) is not easy to deal with, mainly because of the lack of estimates on the process Z (see Pardoux, 1999, Section 6, p. 535 for a discussion).

In this article, we deal with a process that is generated by a divergence form operator L which may also be written as a non-divergence form operator. So, one could use a weak solution u of (4) to define some BSDE of type (5) by identifying (Y_t, Z_t) with $(u(t, X_t), \nabla u(t, X_t))$. Solving PDEs with divergence form operators requires much less regularity on the coefficients than for PDEs with non-divergence form operators. With a bit of regularity on the diffusion coefficient, it is then possible to transform L into a non-divergence form operator, so that X is also a weak solution of some SDE which involves $u(t, X_t)$ and $\nabla u(t, X_t)$. Substituting Y_t and Z_t to $u(t, X_t)$ and $\nabla u(t, X_t)$ allows to conclude. We have to note however that, when one transforms quasi-linear differential divergence form operators into non-divergence form operators, a term which is quadratic in ∇u appears. But the PDE which is involved may still be solved in a way that the solution remains in the space in which we have developed our results on BSDEs. With this method, no regularity is required on $(t, x) \mapsto b(t, x, \cdot, \cdot)$ and to $(t, x) \mapsto h(t, x, \cdot, \cdot)$.

Notations

The set $\mathcal{O}$ will be either $\mathbb{R}^N$ or a bounded, open set of $\mathbb{R}^N$ with a smooth boundary The Euclidean norm in $\mathbb{R}^N$ is denoted by $\|\cdot\|$.

Throughout this article, we use the standard notations about functional spaces. So, for $1 \leq p, q \leq \infty$, $\mathrm{L}^{q,p}(s, T; \mathcal{O})$ denotes the space of measurable functions on $[s, T] \times \mathcal{O}$ such that $\|f\|_{\mathrm{L}^{q,p}(s,T;\mathcal{O})} = \left(\int_s^T \left(\int_{\mathcal{O}} |f(t, x)|^p \,\mathrm{d}x\right)^{q/p} \mathrm{d}t\right)^{1/q}$ is finite. The notation $\mathrm{W}^{1,p}(\mathcal{O})$ denotes the space of functions in $\mathrm{L}^p(\mathcal{O})$ with weak derivatives in $\mathrm{L}^p(\mathcal{O})$, equipped with its usual norm: $\|f\|_{\mathrm{W}^{1,p}(\mathcal{O})} = \|f\|_{\mathrm{L}^p(\mathcal{O})} + \sum_{i=1}^N \|\partial_{x_i} f\|_{\mathrm{L}^p(\mathcal{O})}$. The space $\mathrm{H}_0^1(\mathcal{O})$ is the completion of the space of smooth functions with compact support on $\mathcal{O}$ with respect to the norm $\mathrm{W}^{1,2}(\mathcal{O})$. Finally, for a Banach space X, we denote by $\mathrm{L}^p(s, T; \mathrm{X})$ the space of functions f on $[s, T]$ with values in X, and such that $\int_s^T \|f\|_{\mathrm{X}}^p$ is finite.

Besides, we say that $g = (g_1, \dots, g_N)$ belongs to $\mathrm{L}^{q,p}(s, T; \mathbb{R}^N)$ if g_i belongs to $\mathrm{L}^{q,p}(s, T; \mathbb{R}^N)$ for $i = 1, \dots, N$, and we denote by $\|g\|_{\mathrm{L}^{q,p}(s,T;\mathbb{R}^N)}$ the norm $\left(\sum_{i=1}^N \|g_i\|^2_{\mathrm{L}^{q,p}(s,T;\mathbb{R}^N)}\right)^{1/2}$.

Given a stochastic process Y and a real s, we denote by $\mathcal{F}_s^Y = (\mathcal{F}_{s,t}^Y)_{t \geq s}$ the smallest filtration containing $\sigma(Y_r; r \in [s, t])$ and satisfying the usual hypotheses. When the reference to s is not ambiguous, we write $\mathcal{F}^Y$ instead of $\mathcal{F}^Y_s$.

1 On processes generated by time-inhomogeneous divergence form operators

1.1 Parabolic PDEs and fundamental solutions

By a divergence form operator, we mean an operator of type (1), where the coefficients a and b satisfy, for some positive constants λ and Λ,

- $\star$ a and b are measurable on $[0,T]\times\mathbb{R}^N$, (7a)
- $\star$ $a(t,x)=(a_{i,j}(t,x))_{i,j=1,\dots,N}$ is a symmetric $N\times N$-matrix, (7b)
- $\star$ $\lambda|\xi|^2\leq a_{i,j}(t,x)\xi_i\xi_j,\ \forall\xi\in\mathbb{R}^N,\ \forall(t,x)\in[0,T]\times\mathbb{R}^N$, (7c)
- $\star$ $b(t,x)$ is a vector with values in $\mathbb{R}^N$, (7d)
- $\star$ $|a_{i,j}(t,x)|\leq\Lambda,\ |b_i(t,x)|\leq\Lambda,\ \forall i,j=1,\dots,N,\ \forall(t,x)\in[0,T]\times\mathbb{R}^N$, (7e)

Throughout all this article, we use the convention that $C_1,C_2,\dots$ denote some positive constants that depend only on λ, Λ, the dimension N and T and some given positive reals p and q.

By a solution of the PDE

$$\frac{\partial u}{\partial t}(t,x)+Lu(t,x)=f(t,x)+\frac{\partial g_i}{\partial x_i}(t,x)\text{ with }u(T,x)=h(x),\tag{8}$$

we mean an element of $\mathrm{L}^2(0,T;\mathrm{H}^1_0(\mathbb{R}^N))$ which is a *weak solution*, *i.e.*, for any smooth function φ on $[0,T]\times\mathbb{R}^N$ with $\varphi(0,\cdot)=0$,

$$\begin{aligned}\int_0^T\int_{\mathbb{R}^N}\frac{\partial\varphi}{\partial t}u(t,x)\,\mathrm{d}t\,\mathrm{d}x+\frac{1}{2}\int_0^T\int_{\mathbb{R}^N}a\nabla u\nabla\varphi(t,x)\,\mathrm{d}t\,\mathrm{d}x\\ -\int_0^T\int_{\mathbb{R}^N}b\nabla u\varphi(t,x)\,\mathrm{d}t\,\mathrm{d}x=-\int_0^T\int_{\mathbb{R}^N}f\varphi(t,x)\,\mathrm{d}t\,\mathrm{d}x\\ +\int_0^T\int_{\mathbb{R}^N}g\nabla\varphi(t,x)\,\mathrm{d}t\,\mathrm{d}x+\int_{\mathbb{R}^N}\varphi(T,x)h(x)\,\mathrm{d}x.\end{aligned}$$

In fact, a version of u belongs to $\mathcal{W}_{0,T}$, where for any $s\in[0,T)$, $\mathcal{W}_{s,T}=\mathcal{C}(s,T;\mathrm{L}^2(\mathbb{R}^N))\cap\mathrm{L}^2(s,T;\mathrm{H}^1_0(\mathbb{R}^N))$. This means that $t\mapsto u(t,\cdot)$ is continuous from $[0,T)$ to $\mathrm{L}^2(\mathbb{R}^N)$. This space $\mathcal{W}_{s,T}$ is equipped with the norm

$$\|u\|_{\mathcal{W}_{s,T}}=\left(\sup_{s\leq t\leq T}\|u(t,\cdot)\|^2_{\mathrm{L}^2(\mathbb{R}^N)}+\int_s^T\|\nabla u(t,\cdot)\|^2_{\mathrm{L}^2(\mathbb{R}^N)}\,\mathrm{d}t\right)^{1/2}.$$

The PDE (8) can be solved if h belongs to $\mathrm{L}^2(\mathbb{R}^N)$ and if f and $g_1, \dots, g_N$ belong to $\mathrm{L}^2(0,T;\mathrm{L}^2(\mathbb{R}^N))$. The norm of u in $\mathcal{W}_{s,T}$ may be estimated from that of g, f, h, the bounds of the coefficients and the constant of uniform ellipticity:

$$\|u\|^2_{\mathcal{W}_{s,T}} \leq C_1 \|h\|^2_{\mathrm{L}^2(\mathbb{R}^N)} + C_2 \|f\|^2_{\mathrm{L}^{2,2}(s,T;\mathbb{R}^N)} + C_3 \sum_{i=1}^N \|g_i\|^2_{\mathrm{L}^{2,2}(s,T;\mathbb{R}^N)}. \tag{9}$$

This inequality is called the *energy inequality* (See for example Theorem 3 in Aronson, 1968, p. 639 or Ladyženskaja et al., 1968, Chapter III).

Besides, Equation (8) can also be solved if f belongs to $\mathrm{L}^{q,p}(0,T;\mathbb{R}^N)$ and $g_1, \dots, g_N$ belong to $\mathrm{L}^{m,r}(0,T;\mathbb{R}^N)$, where p, q, r, m satisfy

$$2 < p, q \leq \infty,\ \frac{N}{2p} + \frac{1}{q} < 1, \tag{10a}$$

$$\text{and } 1 < r, m \leq \infty,\ \frac{N}{2r} + \frac{1}{m} < \frac{1}{2}. \tag{10b}$$

In this case, there exists a version of u which is continuous on each compact subset of $[0,T) \times \mathbb{R}^N$: See Aronson, 1968. When a continuous version of u exists, then u denotes in fact this continuous version.

It was shown in Aronson, 1968 that the operator $\partial_t + L$ has a fundamental solution Γ, that is a function $\Gamma(s,y,t,x)$ such that $\partial_s \Gamma(s,y,t,x) + L_y \Gamma(s,y,t,x) = 0$ for any $(s,y) \in [0,t) \times \mathbb{R}^N$ and $\Gamma(s,y,t,x)$ converges weakly to the Dirac δ_{y-x} as s increases to t. The solution u of (8) on $[s,t]$ with $f = g_1 = \cdots = g_N = 0$ and the final condition $u(t,x) = h(x)$ is given by $u(s,y) = \int_{\mathbb{R}^N} \Gamma(s,y,t,x) h(x)\,\mathrm{d}x$ for any $(s,y) \in [0,t] \times \mathbb{R}^N$.

Among the important features of Γ are that there exist C_1, C_2, C_3 and C_4 such that

$$\frac{C_1}{t^{N/2}} \exp\left(-\frac{C_3|x-y|^2}{t-s}\right) \leq \Gamma(s,y,t,x) \leq \frac{C_3}{t^{N/2}} \exp\left(-\frac{C_4|x-y|^2}{t-s}\right). \tag{11}$$

This estimate (11) is called the *Aronson estimate*. If p', q', r' and m' are respectively the Hölder conjugates of p, q, r and m, where (q,p) satisfy (10a) and (m,r) satisfy (10b), then

$$\|\Gamma(s,y,\cdot,\cdot)\|_{\mathrm{L}^{q',p'}(s,T;\mathbb{R}^N)} \leq C_1 \text{ and } \|\nabla\Gamma(s,y,\cdot,\cdot)\|_{\mathrm{L}^{m',r'}(s,t;\mathbb{R}^N)} \leq C_2, \tag{12}$$

where $\nabla\Gamma(s,y,t,x)$ denotes the derivative with respect to x. Finally, for any $(s,y) \in [s,T) \times \mathbb{R}^N$ and any $\delta > 0$, $\Gamma(s,y,\cdot,\cdot)$ belongs to $\mathrm{L}^2(s+\delta,T;\mathrm{H}^1_0(\mathbb{R}^N))$.

1.2 Existence of a stochastic process

The fundamental solution also satisfies $\int_{\mathbb{R}^N} \Gamma(s, y, t, x)\, \mathrm{d}x = 1$ for any $0 \leq s \leq t \leq T$. With the upper bound of the Aronson estimate (11), it is easily proved that Γ is a transition density function of a continuous, conservative, strong Markov process $(\Omega, \mathcal{F}_\infty, X_t, \mathbb{P}_{s,y}, \mathcal{F}_{s,t}; s \in [0,T], t \in [s,T], y \in \mathbb{R}^N)$. Here, $\mathbb{P}_{s,y}$ is such that $\mathbb{P}_{s,y}[\, X_t = y,\ 0 \leq t \leq s\,] = 1$. For any probability measure ν on $\mathbb{R}^N$, we use the notation $\mathbb{P}_{s,\nu}$ to denote the probability measure $\int \nu(\,\mathrm{d}x)\mathbb{P}_{s,x}$. The filtration $\mathcal{F} = (\mathcal{F}_{s,t})_{t \geq s}$ is the minimal filtration to which $(X_t)_{t \geq s}$ is adapted and complete under $\mathbb{P}_{s,\mu}$ for any measure μ (in particular the filtration $\mathcal{F}$ is right-continuous: See for example Blumenthal and Getoor, 1968 for the construction of the Markov process and the filtration $\mathcal{F}$). By construction, for any Borel set B,

$$\mathbb{P}_{s,y}[\, X_t \in B \,|\, \mathcal{F}_{s,u}\,] = \int_{\mathbb{R}^N} \Gamma(u, X_u, t, z) \mathbf{1}_B(z)\, \mathrm{d}z.$$

One of the practical feature of the Aronson estimate is that

$$\mathbb{P}_{s,y}\left[\sup_{r \in [s,t]} |X_r - y| \geq R\right] \leq C_1 \exp\left(\frac{-C_2 R^2}{t-s}\right) \tag{13}$$

for any $R \geq 0$ and any $t \in [s,T]$ (see for example Lemma II.1.2 in Stroock, 1988 in the time-homogeneous case). This estimate (13) is important because it allows to use a localization argument.

Let $\mathcal{O}$ be an open bounded set of $\mathbb{R}^N$ with a smooth boundary. We assume that the coefficients a and b are only defined in $\mathcal{O}$. The PDE (8) has still a solution in $\mathcal{W}_{0,T} = \mathrm{L}^2(0,T;\mathrm{H}^1_0(\mathcal{O})) \cap \mathcal{C}(0,T;\mathrm{L}^2(\mathcal{O}))$. This means that we consider that the solution, if it is smooth, satisfies $u(t,x) = 0$ when $t \in [s,T)$ and x belongs to the boundary of $\mathcal{O}$. Except the lower bound of the Aronson estimate (11), all the results given previously on PDEs on $[0,T] \times \mathbb{R}^N$ (energy estimate, continuity of the solutions, ...) are still true in this case: On that topic, see for example Chapter III in Ladyženskaja et al., 1968. The fundamental solution of (8) is also denoted by Γ and satisfies (12) and the upper bound in the Aronson estimate (11).

When the coefficients of the operator L are defined only on $[0,T] \times \mathcal{O}$, then they are extended on $[0,T] \times \mathbb{R}^N$ by setting $a(t,x) = \lambda \mathrm{Id}$ and $b(t,x) = 0$ for any $x \in \mathbb{R}^N \setminus \mathcal{O}$ and $t \in [0,T]$. Any other function is extended to be zero outside $\mathcal{O}$, so including the boundary of $\mathcal{O}$. This choice is not so arbitrary, since we consider PDEs with a Dirichlet boundary condition equal to 0 on $\partial\mathcal{O}$. Hence, the part of the trajectories of X after the first exit time from $\mathcal{O}$ will not be taken into consideration.

1.3 A convergence result

We assume that $\mathcal{O} = \mathbb{R}^N$ until Section 2.

We call $(a^n, b^n)_{n\in\mathbb{N}}$ a *sequence of smooth approximations of* (a, b) if a^n and b^n are smooth, satisfy (7a)–(7e) with the same constants λ and Λ, and $a^n(t, x)$ and $b^n(t, x)$ converges to $a(t, x)$ and $b(t, x)$ for any t and for almost every x in $\mathbb{R}^N$.

Let X^n be the process generated by L^n, where L^n is the divergence form operator (1), where a and b are replaced by a^n and b^n. Hence, one knows that X^n converges in distribution to X (See for example Rozkosz, 1996b).

Lemma 1. *Let $(s, y) \in [0, T] \times \mathbb{R}^N$ be fixed. Let $(a^n, b^n)_{n\in\mathbb{N}}$ be a sequence of smooth approximations of (a, b), and let M^n be the martingale part of the semi-martingale X^n. Then there exists a square-integrable $\mathcal{F}$-martingale M such that, at least along a subsequence, $\mathbb{P}_{s,y} \circ (X^n, M^n)^{-1}$ converges in distribution to $\mathbb{P}_{s,y} \circ (X, M)^{-1}$.*

Remark 1. At this point, nothing allows to assert that the limit (X, M) is unique nor it is independent from the choice of $(a^n, b^n)_{n\in\mathbb{N}}$. However, we will see in Theorem 1 that the limit is unique and does not depend on the sequence $(a^n, b^n)_{n\in\mathbb{N}}$.

Let $(\nu^n)_{n\in\mathbb{N}}$ be a sequence of probability distribution, such that ν^n converges in distribution to ν. Let also (f^n) and (g^n) be some sequence of functions. Let $\delta \geq 0$. We consider three distinct hypotheses:

(H-i) $\delta > 0$, f^n converges to f in $\mathrm{L}^1([s + \delta, T] \times \mathbb{R}^N)$ and g^n converges to g in $\mathrm{L}^2([s + \delta, T] \times \mathbb{R}^N)$.

(H-ii) $\delta = 0$, f^n converges to f in $\mathrm{L}^1([s, T] \times \mathbb{R}^N)$, g^n converges to g in $\mathrm{L}^2([s, T] \times \mathbb{R}^N)$ and ν^n has a density which is bounded uniformly in n.

(H-iii) $\delta = 0$, f^n converges to f in $\mathrm{L}^{q,p}(s, T; \mathbb{R}^N)$ and g^n converges to g in $\mathrm{L}^{m,r}(s, T; \mathbb{R}^N)$, where (q, p) satisfies (10a) and (m, r) satisfies (10b).

The proof of the next proposition relies on some standard arguments: See for example Rozkosz, 1996a; Lejay, 2002a, ...

Proposition 1. *Let s be fixed. Let $(\nu^n)_{n\in\mathbb{N}}$ be a sequence of probability measures. Let $(M^n)_{n\in\mathbb{N}}$ be a sequence of square-integrable $\mathcal{F}^{X^n}$-martingales on $[s, T]$ such that*

$$\sup_{n\in\mathbb{N}} \sup_{i=1,\dots,N} \mathbb{E}_{\nu^n}[\,\langle M^{i,n}\rangle_T\,] < +\infty.$$

We assume that $\mathbb{P}_{\nu^n} \circ (X^n, M^n)^{-1}$ *converges to* $\mathbb{P} \circ (X, M)^{-1}$ *for some distribution* $\mathbb{P}$*. Then,* M *is a square-integrable* $\mathcal{F}^{X,M}$*-martingale and, under either* (H-i)*,* (H-ii) *or* (H-iii)*,*

$$\mathbb{P}_{s,\nu^n} \circ \left(X^n, M^n, \int_{s+\delta}^{\cdot} f^n(r, X^n_r)\,\mathrm{d}r, \int_{s+\delta}^{\cdot} g^n(r, X^n_r)\,\mathrm{d}M^n_r \right)^{-1} \\ \xrightarrow[n\to\infty]{} \mathbb{P} \circ \left(X, M, \int_{s+\delta}^{\cdot} f(r, X_r)\,\mathrm{d}r, \int_{s+\delta}^{\cdot} g(r, X_r)\,\mathrm{d}M_r \right)^{-1}$$

in the space of continuous functions on $[s+\delta, T]$*. Moreover,* $\mathbb{P} \circ X^{-1}$ *is* $\mathbb{P}_{s,\nu}$ *and if* M *is* $\mathcal{F}^X$*-adapted, then* M *is a square-integrable* $(\mathbb{P}_{s,\nu}, \mathcal{F})$*-martingale.*

Proof of Lemma 1. Since a^n is uniformly bounded, $(\langle M^n \rangle)_{n\in\mathbb{N}}$ is tight, and so is $(M^n)_{n\in\mathbb{N}}$. If (X, M) denotes the limit of a subsequence of $(X^n, M^n)_{n\in\mathbb{N}}$, M is a square-integrable $\mathcal{F}^{X,M}$-martingale. Thanks to Proposition 1, its cross-variations are $\langle M^i, M^j \rangle_t = \int_s^t a_{i,j}(r, X_r)\,\mathrm{d}r$, $\mathbb{P}_{s,y}$-almost surely. We have to note that *a priori*, the distribution of (X, M) is an extension of $\mathbb{P}_{s,y}$, which we still denote by $\mathbb{P}_{s,y}$. Without loss of generality, we assume that the whole sequence $(X^n, M^n)_{n\in\mathbb{N}}$ converges in distribution to (X, M), and not only a subsequence.

Let f be a smooth function with compact support. and u^n be the solution of $\partial_t u^n(t,x) + L^n u^n(t,x) = f(t,x)$ on $[s,T] \times \mathbb{R}^N$ with a given final condition $u(T,x) = \psi(x)$, where ψ is a smooth function in $\mathrm{L}^2(\mathbb{R}^N)$. Let us fix a positive δ. According to the Itô formula, for any $t \geq s+\delta$,

$$u^n(t, X^n_t) = u^n(s+\delta, X^n_{s+\delta}) + \int_{s+\delta}^t f(r, X^n_r)\,\mathrm{d}r + \int_{s+\delta}^t \nabla u^n(r, X^n_r)\,\mathrm{d}M^n_r. \quad (14)$$

Let u be the weak solution of $\partial_t u(t,x) + Lu(t,x) = f(t,x)$ on $[s,T] \times \mathbb{R}^N$ and the final condition $u(T,x) = \psi(x)$. This solution u belongs to $\mathcal{W}_{s,T}$ and $\|u - u^n\|_{\mathcal{W}_{s,T}}$ converges to 0 (See for example Theorem III.4.5 in Ladyženskaja et al., 1968, p. 166). For any $t \geq s+\delta$, the Aronson estimate (11) yields

$$\mathbb{E}_{s,y}\left[\, |u^n(t, X^n_t) - u(t, X^n_t)| \,\right] \leq \int_{\mathbb{R}^N} \Gamma(s,y,t,x) |u^n(t,x) - u(t,x)|\,\mathrm{d}x \\ \leq \frac{C_1}{\delta^{N/2}} \|u^n(t,\cdot) - u(t,\cdot)\|_{\mathrm{L}^2(\mathbb{R}^N)}.$$

Since ψ is smooth, u is continuous on $[s,T] \times \mathbb{R}^N$ (See for example Theorem III.7.1 and Theorem III.10.1 in Ladyženskaja et al., 1968, p. 181 and p. 204). So, $u(t, X^n_t)$ converges in distribution to $u(t, X_t)$, and then $u^n(t, X^n_t)$ converges in distribution to $u(t, X_t)$ for any $t \geq s+\delta$. On the other hand, the right-hand side of (14) converges in the space of continuous functions on $[s+\delta,T]$ to $t\mapsto u(s+\delta,X_{s+\delta})+\int_{s+\delta}^t f(r,X_r)\,\mathrm{d}r+\int_{s+\delta}^t \nabla u(r,X_r)\,\mathrm{d}M_r$. Again using the continuity of u, it is easily established that $\mathbb{P}_{s,y}$-almost surely, for any $t\in[s+\delta,T]$,

$$u(t,X_t)=u(s+\delta,X_{s+\delta})+\int_{s+\delta}^t f(r,X_r)\,\mathrm{d}r+\int_{s+\delta}^t \nabla u(r,X_r)\,\mathrm{d}M_r.$$

It follows that for any $\delta>0$, $\int_{s+\delta}^{\cdot} \nabla u(r,X_r)\,\mathrm{d}M_r$ is $\mathcal{F}$-adapted.

Now, let ρ be a smooth function with compact support, such that $\rho=1$ on a bounded, open set $\mathcal{Q}$ of $\mathbb{R}^N$. Let us set $g_i^j(t,x)=\frac{1}{2}\rho(x)a_{i,j}(t,x)$ and $f_i(t,x)=\rho(x)b_i(t,x)$ for $i,j=1,\dots,N$. Let u_i be the unique weak solution of $\partial_t u_i(t,x)+Lu_i(t,x)=f_i+\partial_{x_j}g_i^j(t,x)$ with the final condition $u_i(T,x)=\rho(x)x_i$. Let $(g_i^{j,n},f_i^n)$ be a sequence of smooth functions converging to (g_i^j,f_i) in $\mathrm{L}^2([s,T]\times\mathbb{R}^N)^2$. Let u_i^n be the solution of $\partial_t u_i^n+Lu_i^n=\partial_{x_j}g_i^{j,n}+f_i^n$ with the final condition $u_i^n(T,x)=x_i\rho(x)$. Again, $\|u-u^n\|_{\mathcal{W}_{s,T}}$ converges to 0, and then $N^n=\int_{s+\delta}^{\cdot}\nabla u_i^n(r,X_r)\,\mathrm{d}M_r$ converges in the space of continuous functions on $[s+\delta,T]$ in $\mathrm{L}^2(\mathbb{P}_{s,y})$ to $N=\int_{s+\delta}^{\cdot}\nabla u_i(r,X_r)\,\mathrm{d}M_r$, which is then a $\mathcal{F}$-martingale, since it is $\mathcal{F}_t$-adapted. For that, the Jensen inequality for conditional expectation implies that for any $s\leq r\leq t\leq T$,

$$\mathbb{E}_{s,y}\left[(\mathbb{E}_{s,y}[N_t^n-N_t\,|\,\mathcal{F}_r])^2\right]\leq\mathbb{E}_{s,y}\left[(N_t^n-N_t)^2\right]\xrightarrow[n\to\infty]{}0.$$

So, $\mathbb{E}_{s,y}[N_t^n\,|\,\mathcal{F}_r]=N_r^n$ converges in $\mathrm{L}^2(\mathbb{P}_{s,y})$ to $\mathbb{E}_{s,y}[N_t\,|\,\mathcal{F}_r]$ which is equal to N_r.

The functions g_i^j and f_i are such that $u_i(r,x)=x_i$ and $\nabla u_i(r,x)=e_i$ if $x\in\mathcal{Q}$, where $(e_1,\dots,e_N)$ is the canonical basis of $\mathbb{R}^N$. Thus, $\int_{s+\delta}^t\nabla u_i(r,X_r)\,\mathrm{d}M_r=M_t^i-M_{s+\delta}^i$ on $\{\,t<\tau\,\}$, where τ is the first exit time from $\mathcal{Q}$. As

$$M_{t\wedge\tau}^i=(M_{t\wedge\tau}^i-M_{(s+\delta)\wedge\tau}^i)+(M_{(s+\delta)\wedge\tau}^i-M_s^i)$$

and $\mathbb{E}_{s,y}\left[(M_{s+\delta}^i-M_s^i)^2\right]\leq C_1\delta$, $M_{t\wedge\tau}^i$ is the limit in $\mathrm{L}^2(\mathbb{P}_{s,y})$ of $\mathcal{F}_t$-measurable random variables, and is itself $\mathcal{F}_t$-measurable.

By localization, and since $\mathbb{E}_{s,y}\left[\sup_{t\in[s,T]}|M_t|^2\right]$ is finite, M is in fact a square-integrable $\mathcal{F}$-martingale. A consequence is that M may defined on the probability space $(\Omega,\mathcal{F}_\infty,\mathbb{P}_{s,y})$ on which X is defined, and not necessarily on an extension of this probability space. □

1.4 Time reversal of a diffusion

For a function $f:[s,T]\times\mathbb{R}^N\to\mathbb{R}$, we set $\overline{f}(t,x)=f(T+s-t,x)$.

Let us fix a point $(s,y)\in[0,T)\times\mathbb{R}^N$. We denote by $\Gamma(t,x)$ the function $\Gamma(s,y,t,x)$. We assume in a first time that a and b are smooth. Let $\overline{L}$ be the divergence form operator

$$\overline{L}=\frac{1}{2\overline{\Gamma}^2(t,x)}\frac{\partial}{\partial x_i}\left(\overline{a}_{i,j}(t,x)\overline{\Gamma}^2(t,x)\frac{\partial}{\partial x_j}\right)-\overline{b}_i(t,x)\frac{\partial}{\partial x_i}$$

defined for $(t,x)\in[s,T)\times\mathbb{R}^N$. This operator $\overline{L}$ may be rewritten

$$\overline{L}=\frac{1}{2}\frac{\partial}{\partial x_i}\left(\overline{a}_{i,j}(t,x)\frac{\partial}{\partial x_j}\right)+\left(\frac{\overline{a}_{i,j}(t,x)}{\overline{\Gamma}(t,x)}\frac{\partial\overline{\Gamma}(t,x)}{\partial x_j}-\overline{b}_i\right)\frac{\partial}{\partial x_i}.$$

If a and b are smooth, it follows from the results in Haussmann and Pardoux, 1986 that $\overline{X}=X_{T+s-\cdot}$ defined on $[s,T]$ is a diffusion process whose infinitesimal generator is $\overline{L}$. We have to remember that the initial distribution of $\overline{X}$ is $\Gamma(s,y,T,x)\,\mathrm{d}x$, and it is conditioned to be at y at time T. Of course, $\overline{X}$ is adapted to $(\mathcal{F}^{\overline{X}}_{s,t})_{t\geq s}$, which is the minimal admissible filtration satisfying the usual hypotheses and where $\mathcal{F}^{\overline{X}}_{s,t}$ contains $\sigma(\overline{X}_r;r\leq t)=\sigma(X_r;r\geq T+s-t)$.

Let us denote by $\overline{M}$ the martingale part of $\overline{X}$ under $\mathbb{P}_{s,y}$. For any $\delta>0$, this martingale is a $(\mathcal{F}^{\overline{X}}_{s,t})_{t\in[s,T-\delta]}$-martingale with cross-variations

$$\langle\overline{M}^i,\overline{M}^j\rangle_t=\int_s^t\overline{a}_{i,j}(r,\overline{X}_r)\,\mathrm{d}r=\int_{T+s-t}^T a_{i,j}(r,X_r)\,\mathrm{d}r,\ t\in[s,T-\delta].\quad(15)$$

As a is bounded, one obtains that for $i=1,\dots,N$, $\sup_{t\in[s,T]}\mathbb{E}_{s,y}[\,\langle\overline{M}^i\rangle_t\,]<+\infty$. Hence, from the L^2 theory for martingales, $\overline{M}_T$ is well defined, and $\overline{M}$ is a continuous, square-integrable martingale on $[s,T]$.

Now, let $(a^n,b^n)_{n\in\mathbb{N}}$ be a family of smooth approximations of (a,b). Let $\overline{M}^n$ be the martingale part of the diffusion $\overline{X}^n=X^n_{T+s-\cdot}$. Using the boundedness of a and (15), the sequence $(\langle\overline{M}^n\rangle)_{n\in\mathbb{N}}$ is clearly tight, and so is $(\overline{M}^n)_{n\in\mathbb{N}}$. It follows that $(\mathbb{P}_{s,y}\circ(X^n,M^n,\overline{M}^n)^{-1})_{n\in\mathbb{N}}$ is tight and converges, at least along a subsequence, to $\mathbb{P}_{s,y}\circ(X,M,\overline{M})^{-1}$, where M is one of the possible limits of $(M^n)_{n\in\mathbb{N}}$, and $\overline{M}$ is a continuous process. As $\sup_{n\in\mathbb{N}}\mathbb{E}_{s,y}\left[\sup_{t\in[s,T]}|\overline{M}^n_t|^2\right]<+\infty$, $\overline{M}$ is a square-integrable $\mathcal{F}^{\overline{X},\overline{M}}$-martingale whose cross-variations are also given by (15). It will be proved in Theorem 1 that $\overline{M}$ is also unique and is in fact a $\mathcal{F}^{\overline{X}}$-martingale.

1.5 The decomposition theorem

We consider now a triple $(X,M,\overline{M})$ corresponding to a limit of $(X^n,M^n,\overline{M}^n)$, as defined at the end of the previous section.

For a smooth function $g = (g_1, \dots, g_N)$ with compact support, we set for $s < s_0 \leq s_1 \leq T$,

$$\mathfrak{W}_{s_0,s_1}(g) = \int_{T+s-s_1}^{T+s-s_0} (\overline{a}^{-1}\overline{g})(r, \overline{X}_r)\,\mathrm{d}\overline{M}_r + \int_{s_0}^{s_1} (a^{-1}g)(r, X_r)\,\mathrm{d}M_r,$$
$$\mathfrak{V}_{s_0,s_1}(g) = \int_{s_0}^{s_1} \Gamma^{-1} g \cdot \nabla\Gamma(r, X_r)\,\mathrm{d}r.$$

In fact, we will see below that one could set $s_0 = s$. The main theorem of the Section is the following.

Theorem 1. (i) *The martingales M and $\overline{M}$ are respectively $\mathcal{F}$ and $\mathcal{F}^{\overline{X}}$-martingales and are unique.*

(ii) *Under $\mathbb{P}_{s,y}$ for any $(s,y) \in [0,T] \times \mathbb{R}^N$, the process X is a Dirichlet process with decomposition $X_t = y + M_t + V_t$ with*

$$V_t = -\frac{1}{2}M_t - \frac{1}{2}(\overline{M}_T - \overline{M}_{T+s-t}) - \frac{1}{2}\int_s^t \Gamma^{-1} a \cdot \nabla\Gamma(r, X_r)\,\mathrm{d}r + \int_s^t b(r, X_r)\,\mathrm{d}r. \tag{16}$$

(iii) *Let u be a continuous function on on $[s,T] \times \mathbb{R}^N$ such that for some $p > N \wedge 2$,*

$$\sup_{t\in[0,T]} \|\nabla u(t,\cdot)\|_{\mathrm{L}^p(\mathbb{R}^N)} \text{ and } \sup_{t\in[0,T]} \|\partial_t u(t,x)\|_{\mathrm{L}^p(\mathbb{R}^N)} \text{ are finite.}$$

Let (s,y) belongs to $[0,T] \times \mathbb{R}^N$. Then, under $\mathbb{P}_{s,y}$, $t \mapsto u(t, X_t)$ is a Dirichlet process and for $t \in [s,T]$,

$$u(t, X_t) = u(s, X_s) - \mathfrak{W}_{s,t}(g) - \mathfrak{V}_{s,t}(g) + \int_s^t f(r, X_r)\,\mathrm{d}r + \int_s^t \nabla u(r, X_r)\,\mathrm{d}M_r, \tag{17}$$

where for $i = 1, \dots, N$,

$$g_i(t,x) = \frac{1}{2} a_{i,j}(t,x) \frac{\partial u(t,x)}{\partial x_j} \text{ and } f(t,x) = b_j(t,x)\frac{\partial u(t,x)}{\partial x_j} + \frac{\partial u(t,x)}{\partial t}. \tag{18}$$

Remark 2. Using this Theorem, there should be no real difficulty to extend to time-inhomogeneous processes generated by divergence form operators the results relying on some stochastic integrals of type $\int f(X_s)\mathrm{d}X_s$ whose definitions use the decomposition (16) for time-homogeneous processes: See Rozkosz, 1996a; Lyons and Stoica, 1999; Lejay, 2002b for example.

Lemma 2. *For any $s_0 \in [s,T)$, the process $t \in [s_0,T] \mapsto \mathfrak{V}_{s_0,t}(g)$ is a continuous process of finite variation defined for $g \in \mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$ where (q,p) satisfies* (10a). *Moreover, if $(g_n)_{n\in\mathbb{N}}$ converges to g in $\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$, then $t \mapsto \mathfrak{V}_{s,t}(g_n)$ converges in probability to $t \mapsto \mathfrak{V}_{s,t}(g)$ uniformly in the space of continuous functions.*

Proof. Let us remark that

$$\mathbb{E}_{s,y}\left[\int_s^T \left|\Gamma^{-1}g\cdot\nabla\Gamma(r,X_r)\right|\mathrm{d}r\right]\leq\int_{\mathbb{R}^N}\int_s^T \|g(r,x)\|\,\|\nabla\Gamma(r,x)\|\;\mathrm{d}r\,\mathrm{d}x$$
$$\leq\|g\|_{\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)}\,\|\nabla\Gamma\|_{\mathrm{L}^{(2q)',(2p)'}(s,T;\mathbb{R}^N)}\leq C_1\,\|g\|_{\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)}\,.$$

This yields that $\int_{s_0}^{s_1}\Gamma a\cdot\nabla\Gamma(r,X_r)\,\mathrm{d}r$ for any $s\leq s_0\leq s_1\leq T$ is well defined almost surely, that $t\mapsto\mathfrak{V}_{s,t}(g)$ is continuous on $[s,T]$ and that the convergence of g_n to g in $\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$ implies that of $\mathfrak{V}(g_n)$ to $\mathfrak{V}(g)$. □

Lemma 3. *The process $(\mathfrak{W}_{u,v}(g))_{u,v\in[s,T]}$ is defined for g in $\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$ where (q,p) satisfies* (10a) *and $t\mapsto\mathfrak{W}_{s,t}(g)$ is continuous. Moreover, if g_n converges to g in $\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$, then $t\mapsto\mathfrak{W}_{s,t}(g^n)$ converges to $t\mapsto\mathfrak{W}_{s,t}(g)$ uniformly on $[s,T]$ in probability under $\mathbb{P}_{s,y}$ for any $(s,y)\in[0,T]\times\mathbb{R}^N$.*

Proof. We remark that if g belongs to $\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$, then

$$\mathbb{E}_{s,y}\left[\int_s^T\|g(r,X_r)\|^2\;\mathrm{d}r\right]=\int_{\mathbb{R}^N}\int_s^T\Gamma(r,z)\,\|g(r,z)\|^2\;\mathrm{d}r\,\mathrm{d}z$$
$$\leq\|\Gamma\|_{\mathrm{L}^{q',p'}(s,T;\mathbb{R}^N)}\,\|g\|^2_{\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)}\leq C_1\,\|g\|^2_{\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)}\,.\quad(19)$$

Hence, $\int_s^\cdot g(r,X_r)\,\mathrm{d}M_r$ is a square-integrable $(\mathcal{F},\mathbb{P}_{s,y})$-martingale. In addition, since M is continuous, for any $t>s$, $\int_{s+\delta}^t g(r,X_r)\,\mathrm{d}M_r$ converges almost surely to $\int_s^t g(r,X_r)\,\mathrm{d}M_r$.

Similarly,

$$\mathbb{E}_{s,y}\left[\int_s^T\|\overline{g}(r,\overline{X}_r)\|^2\,\mathrm{d}r\right]\leq C_1\,\|g\|^2_{\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)}\,.\qquad(20)$$

So, $\int_s^\cdot\overline{g}(r,X_r)\,\mathrm{d}\overline{M}_r$ is a square-integrable $(\mathcal{F}^{\overline{X}},\mathbb{P}_{s,y})$-martingale and the limit of $\int_s^{T-\delta}\overline{g}(r,\overline{X}_r)\,\mathrm{d}\overline{M}_r$ exists almost surely and is equal to $\int_s^T\overline{g}(r,\overline{X}_r)\,\mathrm{d}\overline{M}_r$.

The continuity of $t\mapsto\mathfrak{W}_{s,t}(g)$ is clear from the continuity of M and $\overline{M}$. From (19) and (20) and the Doob maximal inequality, it is also clear that if g_n converges to g in $\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$, then $\mathfrak{W}(g_n)$ converges to $\mathfrak{W}(g)$ in probability. □

Lemma 4. *For any $g\in\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$ where (q,p) satisfies* (10a)*, $\mathfrak{W}_{s,t}(g)$ has zero quadratic variation.*

Proof. We assume that $p<\infty$ and $q<\infty$. Let $(g^n)_{n\in\mathbb{N}}$ be a family of smooth functions converging to g in $\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$. Hence, if $s\leq t_1\leq\cdots\leq t_k\leq t$ is a partition in $[s,T]$,

$$\sum_{i=1}^{k-1}\mathbb{E}_{s,y}\left[\mathfrak{W}_{t_i,t_{i+1}}(g)^2\right]\leq\sum_{i=1}^{k-1}\mathbb{E}_{s,y}\left[\mathfrak{W}_{t_i,t_{i+1}}(g-g^n)^2\right]+\sum_{i=1}^{k-1}\mathbb{E}_{s,y}\left[\mathfrak{W}_{t_i,t_{i+1}}(g^n)^2\right].$$

But we remark that for any h in $\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$,

$$\mathbb{E}_{s,y}\left[\mathfrak{W}_{u,v}(h)^2\right] \leq C_1\mathbb{E}_{s,y}\left[\int_u^v \|h\|^2(r,X_r)\,\mathrm{d}r\right] + C_1\mathbb{E}_{s,y}\left[\int_{T+s-v}^{T+s-u} \|h\|^2(r,X_r)\,\mathrm{d}r\right],$$

so that

$$\sum_{i=1}^{k-1}\mathbb{E}_{s,y}\left[\mathfrak{W}_{t_i,t_{i+1}}(g-g^n)^2\right] \leq 2C_1\mathbb{E}_{s,y}\left[\int_s^T \|g-g^n\|^2(r,X_r)\,\mathrm{d}r\right] \\ \leq C_2\|g-g^n\|^2_{\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)}.$$

Using (21), we know that $\mathfrak{W}(g^n)$ is a process of integrable variation, hence of zero quadratic variation. This proves the Lemma.

If $p=\infty$ or $q=\infty$, then let g^ℓ be a function in $\mathrm{L}^{2q,2p}(0,T;\mathbb{R}^N)$ with compact support such that $g^\ell(t,x)=g(t,x)$ on $[s,T]\times B(y,\ell)$ and $\|g-g^\ell\|_{\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)} \xrightarrow[\ell\to\infty]{} 0$, where $B(y,\ell)$ is the ball centered on y and with radius ℓ for some integer ℓ. Let τ^ℓ be the first exit time of X from this ball. For any $C>0$,

$$\mathbb{P}_{s,y}\left[\sum_{i=1}^{k-1}\mathfrak{W}_{t_i,t_{i+1}}(g)^2 > C\right] \leq \mathbb{P}_{s,y}\left[\sum_{i=1}^{k-1}\mathfrak{W}_{t_i,t_{i+1}}(g^\ell)^2 > C\right] + \mathbb{P}_{s,y}\left[\tau^\ell < T\right].$$

With (13), $\mathbb{P}_{s,y}\left[\tau^\ell<T\right]$ decreases to 0 with ℓ. The previous argument on g^ℓ proves that fact that $\mathfrak{W}(g)$ has zero quadratic variation. □

Proposition 2. *Let $g=(g_1,\dots,g_N)$ be a function in $\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$ and such that $\operatorname{div} g$ belongs to $\mathrm{L}^{q,p}(s,T;\mathbb{R}^N)$ for (q,p) satisfying* (10a). *Then, $\mathbb{P}_{s,y}$-almost surely,*

$$\mathfrak{W}_{s_0,s_1}(g) + \mathfrak{V}_{s_0,s_1}(g) = -\int_{s_0}^{s_1}\operatorname{div} g(r,X_r)\,\mathrm{d}r, \tag{21}$$

for any $s\leq s_0\leq s_1\leq T$.

Proof. We assume that a and b are smooth, and that g is also smooth and has compact support. Using the Itô formula both for $\overline{X}$ and X, one remarks that for any $s<s_0\leq s_1\leq T$,

$$\overline{M}_{T+s-s_0} - \overline{M}_{T+s-s_1} = M_{s_0} - M_{s_1} - \int_{s_0}^{s_1}(\Gamma^{-1}a\cdot\nabla\Gamma + \nabla a)(r,X_r)\,\mathrm{d}r. \tag{22}$$

Let $h=(h_1,\dots,h_N)$ be a smooth function with compact support with values in $\mathbb{R}^N$. By definition, $\int_{T+s-s_1}^{T+s-s_0}\overline{h}(r,\overline{X}_r)\,\mathrm{d}\overline{M}_r$ is the limit in probability of $\sum_{i=1}^{k-1}\overline{h}(t_i,\overline{X}_{t_i})(\overline{M}_{t_{i+1}}-\overline{M}_{t_i})$ when the mesh of the partition $T+s-s_1\leq t_1\leq\cdots\leq t_k\leq T+s-s_0$ decreases to 0. Hence, it is easily proved that

$$\begin{aligned}\int_{T+s-s_1}^{T+s-s_0}\overline{h}(r,\overline{X}_r)\,\mathrm{d}\overline{M}_r&=-\int_{s_0}^{s_1}h(r,X_r)\,\mathrm{d}M_r\\&\quad-\int_{s_0}^{s_1}[h_i(\Gamma^{-1}a_{i,j}\partial_{x_j}\Gamma)+a_{i,j}\partial x_j h_i+h_i\partial_{x_j}a_{i,j}](r,X_r)\,\mathrm{d}r.\quad(23)\end{aligned}$$

We note that $\operatorname{div}(ah)=\sum_{i,j=1}^N(a_{i,j}\partial_{x_j}h_i+h_i\partial_{x_j}a_{i,j})$. Thus, if $g=ah$, (23) becomes (21). Up to now, we have assumed that $s_0>s$. But in fact, (21) also holds if $s_0=s$ as the previous Lemmas on $\mathfrak{V}$ and $\mathfrak{W}$ prove it.

If a and b are not smooth, we use a sequence of smooth approximations $(a^n,b^n)_{n\in\mathbb{N}}$ of (a,b). Proposition 1 is easily extended to take into account the convergence of $\overline{M}^n$, the martingale part of $\overline{X}^n$. So, $\mathfrak{W}^n_{s,\cdot}(g)$ converges in distribution to $\mathfrak{W}_{s,\cdot}(g)$, where $\mathfrak{W}^n$ is defined with respect to X^n instead of X. A similar result holds for $\mathfrak{V}$. Finally, using the hypotheses on g and Proposition 1, $\int\operatorname{div}g(r,X^n_r)\,\mathrm{d}r$ converges in distribution to $\int\operatorname{div}g(r,X_r)\,\mathrm{d}r$ jointly with the other convergences. So, (21) holds also when a and b are not smooth. Finally, again using Proposition 1, there is no difficulty to prove that (21) is true under the assumption that g is only weakly differentiable. $\square$

Proof of Theorem 1. We prove first (iii), then (ii) and (i).

Proof of (iii). We assume in a first time that u has compact support. It is clear that u is a weak solution to

$$\frac{\partial u(t,x)}{\partial t}+Lu(t,x)=\operatorname{div}g(t,x)+f(t,x).$$

Due to the hypotheses on u, both f and g_i belong to $\mathrm{L}^{q,p}(s,T;\mathbb{R}^N)$ with $p,q<\infty$. Let $(f^n)_{n\in\mathbb{N}}$ and $(g^n)_{n\in\mathbb{N}}$ be some smooth approximations of f and g. Let u^n be the weak solution of $\partial_t u^n+Lu^n=\operatorname{div}g^n+f^n$ on $[s,T]\times\mathbb{R}^N$ with the final condition $u^n(T,x)=u(T,x)$. For any integer n and any $\delta>0$, the Itô formula, which has already been proved in the proof of Lemma 1, yields that $\mathbb{P}_{s,y}$-almost surely,

$$u^n(t,X_t)=u^n(s+\delta,X_{s+\delta})+\int_{s+\delta}^t(\operatorname{div}g^n+f^n)(r,X_r)\,\mathrm{d}r+\int_{s+\delta}^t\nabla u^n(r,X_r)\,\mathrm{d}M_r,$$

for any $t\in[s+\delta,T]$. With the help of (21), for any $\delta>0$,

$$\begin{aligned}u^n(t,X_t)&=u^n(s+\delta,X_{s+\delta})-\mathfrak{V}_{s+\delta,t}(g^n)-\mathfrak{W}_{s+\delta,t}(g^n)\\&\quad+\int_{s+\delta}^t f^n(r,X_r)\,\mathrm{d}r+\int_{s+\delta}^t\nabla u^n(r,X_r)\,\mathrm{d}M_r.\quad(24)\end{aligned}$$

We have seen that $\mathfrak{W}_{s,\cdot}(g_n)$ and $\mathfrak{V}_{s,\cdot}(g_n)$ converges to $\mathfrak{W}_{s,\cdot}(g)$ and $\mathfrak{V}_{s,\cdot}(g)$ in probability. Since u has a compact support, f and g also belong to $\mathrm{L}^{2,2}(s,T;\mathbb{R}^N)$. Hence, it follows from standard results that $\|u-u^n\|_{\mathcal{W}_{s,T}}$ converges to 0. Using Proposition 1 and the same arguments as in the proof of Lemma 1, $\mathbb{P}_{s,y}$-almost surely,

$$u(t,X_t)=u(s+\delta,X_{s+\delta})-\mathfrak{V}_{s+\delta,t}(g)-\mathfrak{W}_{s+\delta,t}(g)\\ +\int_{s+\delta}^t f(r,X_r)\,\mathrm{d}r+\int_{s+\delta}^t \nabla u(r,X_r)\,\mathrm{d}M_r. \quad (25)$$

for any $t\in[s+\delta,T]$, As u is continuous on $[s,T)\times\mathbb{R}^N$, $u(s+\delta,X_{s+\delta})$ converges to $u(s,y)$ as $\delta\to 0$. We also know that $\int_{s+\delta}^t f(r,X_r)\,\mathrm{d}r$, $\mathfrak{V}_{s+\delta,t}(g)$, $\mathfrak{W}_{s+\delta,t}(g)$ converge almost surely to $\mathfrak{V}_{s,t}(g)$ and $\mathfrak{W}_{s,t}(g)$ as δ decreases to 0.

Finally, $\mathbb{E}_{s,y}\left[\int_s^T \|\nabla u\|^2(r,X_r)\,\mathrm{d}r\right]$ is finite, so that $\int_s^\cdot \nabla u(r,X_r)\,\mathrm{d}M_r$ is well defined as a square-integrable $\mathbb{P}_{s,y}$-martingale. It follows that $\int_{s+\delta}^t \nabla u(r,X_r)\,\mathrm{d}r$ converges almost surely to $\int_s^t \nabla u(r,X_r)\,\mathrm{d}M_t$ for any $t\in[s,T]$. It means that (25) is valid even for $\delta=0$.

Now, if u has not compact support, then let $(\mathcal{O}_k)_{k\in\mathbb{N}}$ be a sequence of increasing open sets of $\mathbb{R}^N$ such that $\cup_{k\in\mathbb{N}}\mathcal{O}_k=\mathbb{R}^N$. Let τ^k be the first exit time from $\mathcal{O}_k$. Let also be ρ_k some smooth functions with compact supports and such that $\rho_k(x)=1$ for $x\in\mathcal{O}_k$. We denote by g^k and f^k the functions given by (18) when u is replaced by $\rho_k u$. So, $-\mathfrak{W}_{s,t}(g)-\mathfrak{V}_{s,t}(g)+\int_s^t f(r,X_r)\,\mathrm{d}r+\int_s^t \nabla u(r,X_r)\,\mathrm{d}M_r$ is equal to $-\mathfrak{W}_{s,t}(g^k)-\mathfrak{V}_{s,t}(g^k)+\int_s^t f^k(r,X_r)\,\mathrm{d}r+\int_s^t \nabla(\rho_k u)(r,X_r)\,\mathrm{d}M_r=u(t,X_t)-u(s,y)$ on $\left\{T\leq\tau^k\right\}$. As τ^k increases almost surely to infinity with k, (17) is true even if u has not compact support.

Proof of (ii). The result is clear with $u(x)=x_i$ for $i=1,\dots,N$, whose derivatives belong to $\mathrm{L}^\infty(\mathbb{R}^N)$.

Proof of (i). We already know that M is a $\mathcal{F}$-martingale. In fact, M is unique since $X_t=y+M_t+V_t$ is a Dirichlet process and this decomposition is unique.

Using a smooth sequence of approximations of (a,b) and the Itô formula, for any smooth function u such that ∇u belongs to $\mathrm{L}^2(\mathbb{R}^N)$, $\int_t^T \nabla u(X_r)\,\mathrm{d}M_r$ is $\mathcal{F}^{\overline{X}}_{T+s-t}$-measurable when $t\geq s$. Using a proof similar to that of Lemma 1, but between times t and T and not $s+\delta$ and T, one obtains easily that for any u in $\mathrm{H}^1_0(\mathbb{R}^N)$, $\int_t^T \nabla u(X_r)\,\mathrm{d}M_r$ is $\mathcal{F}^{\overline{X}}_{T+s-t}$-measurable.

Let $u(t,x)=u(x)$ be a smooth function with compact support, such that $u(x)=x_i$ on the open ball $B(y,k)$ with center y and radius k. Equality (17) written between $T+s-t$ and T yields that

$$\frac{1}{2}\int_s^t \nabla u(\overline{X}_r)\,\mathrm{d}\overline{M}_r = -\frac{1}{2}\int_{T+s-t}^T \nabla u(X_r)\,\mathrm{d}M_r - \int_{T+s-t}^T b\cdot\nabla u(r,X_r)\,\mathrm{d}r$$
$$+\frac{1}{2}\int_{T+s-t}^T \Gamma^{-1}a\nabla u\cdot\nabla\Gamma(r,X_r)\,\mathrm{d}r + u(T,X_T) - u(T+s-t,X_{T+s-t}).$$

It follows that $\int_s^\cdot \nabla u(\overline{X}_r)\,\mathrm{d}\overline{M}_r$ is $\mathcal{F}^{\overline{X}}$-adapted. Let $\overline{\tau}^k$ be the first exit time from $B(y,k)$ for $\overline{X}$. It is a $\mathcal{F}^{\overline{X}}$-stopping time. Thus, $M_{t\wedge\overline{\tau}^k}$ is a square-integrable $\mathcal{F}^{\overline{X}}$-martingale. But

$$\mathbb{P}_{s,y}\left[\overline{\tau}^k < t\right] \leq \mathbb{P}_{s,y}\left[\overline{\tau}^k < T\right] \leq \mathbb{P}_{s,y}\left[\overline{X}\text{ remains in }B(y,k)\text{ on [s,T]}\right]$$

But the later probability is also equal to $\mathbb{P}_{s,y}\left[\,X\text{ remains in }B(y,k)\text{ on [s,T]}\,\right]$, which is known, according to (13), to decrease to 0 as k increases to infinity. As $\mathbb{E}_{s,y}\left[\,\sup_{s\leq t\leq T}|\overline{M}_t|^2\,\right] < +\infty$, $\overline{M}$ is also a $\mathcal{F}^{\overline{X}}$-martingale. The uniqueness of the decomposition (16) proves the uniqueness of $\overline{M}$. □

Let us denote by $\mathcal{D}$ the space of weak solutions of $\partial_t u + Lu = f$ and $u(T,x) = 0$ when f is a continuous function with compact support. In fact, for such a function f, there exists a version of u which is continuous on $[0,T]\times\mathbb{R}^N$, and it is this solution we consider. Let us denote by R the application which maps $u\in\mathcal{D}$ from f.

Lemma 5. *For any smooth function u with compact support on $[0,T]\times\mathbb{R}^N$, there exists a sequence of functions $(u^n)_{n\in\mathbb{N}}$ such that u^n belongs to $\mathcal{D}\cap\mathcal{C}([0,T]\times\mathbb{R}^N)$ and $u^n(t,x)$ converges to $u(t,x)$ for any (t,x). Moreover, $\sup_{n\in\mathbb{N}}\sup_{(t,x)}|u^n(t,x)| < +\infty$.*

Proof. Let us set $g_i(t,x) = \frac{1}{2}a_{i,j}\partial_{x_j}u(t,x)$ and $f(t,x) = \partial_t u(t,x) + b_i\partial_{x_i}u$. Then g_i belongs to $\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$ and f belongs to $\mathrm{L}^{q,p}(s,T;\mathbb{R}^N)$. There exists some sequences $(g_i^n)_{n\in\mathbb{N}}$ and $(f^n)_{n\in\mathbb{N}}$ of smooth functions with compact support such that g_i^n converges to g_i in $\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)$ and f^n converges to f in $\mathrm{L}^{q,p}(s,T;\mathbb{R}^N)$. Let us denote by u^n the weak solution of $\partial_t u^n + Lu^n = f^n + \partial_{x_i}g_i^n$ and $u^n(T,x) = 0$. By definition, u^n belongs to $\mathcal{D}$ and is continuous. According to Theorem 5 in Aronson, 1968, p. 656, u^n is equal to

$$u^n(s,y) = \int_s^T\int_{\mathbb{R}^N}\Gamma(s,y,t,x)f^n(t,x)\,\mathrm{d}x\,\mathrm{d}t - \int_s^T\int_{\mathbb{R}^N}\nabla\Gamma(s,y,t,x)\cdot g^n(t,x)\,\mathrm{d}x\,\mathrm{d}t.$$

Of course, a similar representation holds for u. Using the estimates in (12), the result is clear. Moreover, with (12), $|u^n(t,x)|$ is smaller than $C_1(\|g\|_{\mathrm{L}^{2q,2p}(s,T;\mathbb{R}^N)} + \|f\|_{\mathrm{L}^{q,p}(s,T;\mathbb{R}^N)})$ for any (t,x) in $[s,T]\times\mathbb{R}^N$. □

We say that a distribution $\widehat{\mathbb{P}}$ satisfies the martingale problem at (s,y) if for any function $u=(\partial_t+L)^{-1}f$ in $\mathcal{D}$, $M^u_{s,t}=u(t,X_t)-u(s,X_s)-\int_s^t f(r,X_r)\,\mathrm{d}r$ is a $\widehat{\mathbb{P}}$-martingale and $\widehat{\mathbb{P}}[\,X_s=y\,]=1$. We note that the boundedness of u and f implies that $M^u_{s,\cdot}$ is a square-integrable martingale on $[s,T]$.

Lemma 6. *For any $(s,y)\in[0,T)\times\mathbb{R}^N$, the distribution $\mathbb{P}_{s,y}$ is the unique solution to the martingale problem at (s,y).*

Proof. It is clear that $\mathbb{P}_{s,y}$ is a solution of the martingale problem. Let $\widehat{\mathbb{P}}$ be another solution of the martingale problem at (s,y).

The operator $\alpha-\partial_t-L$ is invertible for any $\alpha>0$ on the image of ∂_t+L of $\mathcal{D}$. Moreover, $(\alpha-\partial_t-L)^{-1}=-e^{\alpha t}R(e^{-\alpha t}\cdot)$, where R is the inverse of ∂_t+L on $\mathcal{D}$. So, using the martingale property, it is standard that that if u belongs to $\mathcal{D}$, both $G_\alpha=\int_0^{+\infty}e^{-\alpha t}\mathbb{E}_{s,y}[\,u(t,X_t)\,]\,\mathrm{d}t$ and $\widehat{G}_\alpha=\int_0^{+\infty}e^{-\alpha t}\widehat{\mathbb{E}}[\,u(t,X_t)\,]\,\mathrm{d}t$ are equal to the inverse of $\alpha-\partial_t-L$ and are consequently equal for any $\alpha>0$. So, $\mathbb{E}_{s,y}[\,u(t,X_t)\,]=\widehat{\mathbb{E}}[\,u(t,X_t)\,]$. With Lemma 5, this is also true for any smooth function u with compact support. Hence, $\widehat{\mathbb{P}}\circ X_t^{-1}=\mathbb{P}_{s,y}\circ X_t^{-1}$. The standard proof for the uniqueness of the martingale problem (see for example 5.4.E in Karatzas and Shreve, 1991, p. 325) is easily adapted to prove that $\widehat{\mathbb{P}}=\mathbb{P}_{s,y}$. □

Theorem 2. *A martingale representation theorem holds with respect to the martingale part M of X with respect to $\mathcal{F}^X$ under $\mathbb{P}_{s,y}$ for any $(s,y)\in[0,T]\times\mathbb{R}^N$.*

Proof. This theorem follows from the uniqueness of the solution of the martingale problem: See for example Lejay, 2002a. □

1.6 Application: The Feynman-Kac formula

Let us consider the operator $A=L+c-\partial_{x_i}(d_i\cdot)$, where c and $d_1,\dots,d_N$ are measurable functions on $[0,T]\times\mathbb{R}^N$. We assume that c and d are bounded by Λ. Accordingly, c and d belong to $\mathrm{L}^{q,p}(0,T;\mathbb{R}^N)$ for $p=q=\infty$, and then (p,q) satisfies (10a) and (10b).

Proposition 3. *The linear Feynman-Kac formula* (3) *is valid for any $(s,y)\in[0,T)\times\mathbb{R}^N$ and any $g\in\mathrm{L}^2(\mathbb{R}^N)$, where $(X,\mathbb{P}_{s,y};(s,y)\in[0,T]\times\mathbb{R}^N)$ is the process generated by $L^0=\frac{1}{2}\partial_{x_i}(a_{i,j}\partial_{x_j})$ with a transition density function Γ.*

Proof. Let $(d^n)_{n\in\mathbb{N}}$ be a family of smooth functions. Let also A^n be the operator $A^n=L+c-\frac{\partial}{\partial x_i}(d^n_i\cdot)$. On $[0,T)\times\mathbb{R}^N$, any solution u^n of $\partial_t u^n+$

$A^n u^n(t,x) = 0$ and $u^n(T,x) = g(x)$ for some $t \leq T$ is also solution to the PDE

$$\frac{\partial u^n(t,x)}{\partial t} + Lu^n(t,x) + c(t,x)u^n(t,x) - d_i^n \frac{\partial u^n}{\partial x_i}(t,x) - \operatorname{div}(d^n)u^n(t,x) = 0,$$

with the final condition $u^n(T,x) = g(x) \in \mathrm{L}^2(\mathbb{R}^N)$ on $[0,T) \times \mathbb{R}^N$.

We assume that $(X, \mathbb{P}_{s,y}; (s,y) \in [0,T) \times \mathbb{R}^N)$ is the process generated by $\frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{i,j}\frac{\partial}{\partial x_j}\right)$. Its transition density function is Γ. Let (s,y) be a fixed point in $[0,T) \times \mathbb{R}^N$.

We assume that u^n is the version of u^n given by $u^n(s,y) = \int_{\mathbb{R}^N} \Upsilon^n(s,y,T,x)g(x)\,\mathrm{d}x$, where Υ^n is the transition density function of A^n. There is no difficulty to adapt in our context the results in Chen and Zhao, 1995 (See also Lejay, 2000, Proposition 0.8, p. 40) to prove that $u^n(s,y)$ is given by the following Girsanov theorem and the Feynman-Kac formula:

$$u^n(s,y) = \mathbb{E}_{s,y}\left[\exp(V^n)g(X_T)\right]$$

with

$$V^n = \int_s^T a^{-1}(b-d^n)(r,X_r)\,\mathrm{d}M_r - \frac{1}{2}\int_s^T a^{-1}(b-d^n)(b-d^n)(r,X_r)\,\mathrm{d}r \\ + \int_s^T c(r,X_r)\,\mathrm{d}r - \int_s^T \operatorname{div} d^n(r,X_r)\,\mathrm{d}r. \quad (26)$$

Using (21), one obtains that

$$V^n = N^n + U^n - W^n$$

with

$$U^n = \int_s^T \Gamma^{-1} d^n \cdot \nabla\Gamma(r,X_r)\,\mathrm{d}r + \int_s^T c(r,X_r)\,\mathrm{d}r, \quad (27)$$

$$N^n = \int_s^T a^{-1}b(r,X_r)\,\mathrm{d}M_r + \int_s^T \overline{a}^{-1}\overline{d}^n(r,\overline{X}_r)\,\mathrm{d}\overline{M}_r \quad (28)$$

$$W^n = \frac{1}{2}\int_s^T a^{-1}(b-d^n)(b-d^n)(r,X_r)\,\mathrm{d}r \quad (29)$$

Let us define V, N and W by (26), (28) and (29) by replacing d^n by d.

We assume that $d^n = \rho_n \star d$, where $(\rho_n)_{n\in\mathbb{N}}$ is a family of mollifiers. Hence, $(d^n)_{n\in\mathbb{N}}$ converges to d in the sense that for every bounded, open set $\mathcal{O}$, $\|d^n - d\|_{\mathrm{L}^{q,p}(0,T;\mathcal{O})}$ converges to 0, where (q,p) satisfies (10b). With a localization argument and the proofs of Lemma 2 and Lemma 3, it is clear that V^n converges in probability under $\mathbb{P}_{s,y}$ to V. Moreover, $d^n(t,x)$ converges also to $d(t,x)$ almost everywhere. Hence, we know that $u^n(s,y) = \int_{\mathbb{R}^N} \Upsilon^n(s,y,T,x)g(x)\,\mathrm{d}x$ converges to $u(s,y) = \int_{\mathbb{R}^N} \Upsilon(s,y,T,x)g(x)$, where Υ is the transition density function of A (See for example Aronson, 1968; Rozkosz, 1996b). So, to obtain formula (3), it remains to prove that $\mathbb{E}_{s,y}\left[\exp(V^n)g(X_T)\right]$ converges to $\mathbb{E}_{s,y}\left[\exp(V)g(X_T)\right]$. For that, we have only to prove that $(\exp(V^n)g(X_T))_{n\in\mathbb{N}}$ is uniformly integrable. If we assume in a first time that g is bounded, this is true for example if $\sup_{n\in\mathbb{N}}\mathbb{E}_{s,y}\left[\exp(4V^n)\right]$ is finite.

With the Cauchy-Schwarz formula,

$$\begin{aligned}\mathbb{E}_{s,y}&\left[\exp\left(4N^n - 32W^n + 28W^n + 4U^n\right)\right]\\ &\leq \mathbb{E}_{s,y}\left[\exp\left(8N^n - 64W^n + 8U^n\right)\right]^{1/2}\mathbb{E}_{s,y}\left[\exp(58W^n)\right]^{1/2}.\end{aligned}$$

But W^n is bounded by $\lambda 2\Lambda^2 T$, and the first expectation in the right-hand side of the previous equation is equal to $\int_s^T \Theta^n(s,y,T,x)\,\mathrm{d}x$, where Θ^n is the transition density function of the differential operator

$$\frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{i,j}\frac{\partial}{\partial x_j}\right) + 8b_i\frac{\partial}{\partial x_i} + 8c - 8\frac{\partial}{\partial x_i}(d_i^n\cdot).$$

The Aronson estimate is still true for differential operators with terms of any order. Hence, $\sup_{n\in\mathbb{N}}\int_s^T \Theta^n(s,y,T,x)\,\mathrm{d}x$ is finite.

We assume now that g belongs only to $\mathrm{L}^2(\mathcal{O})$. Let $(g^n)_{n\in\mathbb{N}}$ be a family of smooth functions converging to g in $\mathrm{L}^2(\mathcal{O})$. Hence,

$$\begin{aligned}\mathbb{E}_{s,y}&\left[\exp(V)|g(X_T) - g^n(X_T)|\right]\\ &\leq \mathbb{E}_{s,y}\left[\exp(2V)\right]^{1/2}\left(\int_{\mathbb{R}^N}\Gamma(s,y,T,x)|g-g^n|^2(x)\,\mathrm{d}x\right)^{1/2}.\end{aligned}$$

Since V^n converges to V in probability and $\sup_{n\in\mathbb{N}}\mathbb{E}_{s,y}\left[\exp(4V^n)\right]$ is finite, it is clear that $\mathbb{E}_{s,y}\left[\exp(2V^n)\right]$ converges to $\mathbb{E}_{s,y}\left[\exp(2V)\right]$ which is then finite. Moreover, $\int_{\mathbb{R}^N}\Gamma(s,y,T,x)|g-g^n|^2(x)\,\mathrm{d}x$ converges to 0. The Proposition is then proved. □

2 Non-linear PDEs and BSDEs

2.1 BSDEs and semi-linear PDEs

We are now interested in semi-linear PDEs of type

$$\frac{\partial u}{\partial t}(t,x) + Lu(t,x) + h(t,x,u(t,x),\nabla u(t,x)) = 0 \tag{30}$$

with the final condition $u(T,x) = g(x)$ on the cylinder $[0,T] \times \mathcal{O}$, where $\mathcal{O}$ is either an open, bounded subset of $\mathbb{R}^N$ or $\mathbb{R}^N$. Of course, L is still a divergence form operator, whose coefficients satisfy (7a)–(7e). Moreover, we do the following assumptions on h and g: For any (t,x,y,y',z,z') in $[0,T] \times \mathcal{O} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N$, there exist some constants C and C' such that

$$\star\ (t,x,y,z) \mapsto h(t,x,y,z) \text{ is measurable,} \tag{31a}$$

$$\star\ |h(t,x,y,z) - h(t,x,y',z)| \leq C|y-y'|, \tag{31b}$$

$$\star\ |h(t,x,y,z) - h(t,x,y,z')| \leq C' \|z - z'\|, \tag{31c}$$

$$\star\ h(t,x,0,0) \in \mathrm{L}^{2,2}(0,T;\mathcal{O}), \tag{31d}$$

$$\star\ g \in \mathrm{L}^2(\mathcal{O}). \tag{31e}$$

We remark that with these hypotheses, if u belongs to $\mathcal{W}_{0,T}$, then $(t,x) \mapsto h(t,x,u(t,x),\nabla u(t,x))$ belongs also to $\mathrm{L}^{2,2}(0,T;\mathcal{O})$.

Let $\mathcal{N}(h)$ be the set of points of $[0,T) \times \mathcal{O}$ such that

$$\mathbb{E}_{s,y}\left[\int_s^{\tau\wedge T} |h(t,X_t,0,0)|^2 \,\mathrm{d}t\right] = +\infty,$$

where τ is the first exit time from $\mathcal{O}$ of the process X. The Lebesgue measure of $\mathcal{N}(h)$ is zero, since for any smooth, positive function φ with $\int_{\mathcal{O}} \varphi(x)\,\mathrm{d}x = 1$, the Aronson estimate (11) implies that

$$\int_{\mathcal{O}}\int_{\mathcal{O}}\int_s^T \varphi(y)\Gamma(s,y,t,x)|h(t,x,0,0)|^2 \,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}x \leq C_1 \|h\|^2_{\mathrm{L}^{2,2}(s,T;\mathcal{O})}.$$

Moreover, $\mathcal{N}(h)$ could be empty, since, according to (12), for any f in $\mathrm{L}^{2q,2p}(s,T;\mathcal{O})$ where (q,p) satisfies (10a), $\int_{\mathcal{O}}\int_s^T \Gamma(s,y,t,x)|f(t,x)|^2\,\mathrm{d}x\,\mathrm{d}t \leq C_1 \|f\|^2_{\mathrm{L}^{2q,2p}(s,T;\mathcal{O})}$.

Proposition 4. *There exists a constant C depending only on λ, Λ, T and N such that $\mathcal{N}(h) \subset \mathcal{N}^*(h)$ where $\mathcal{N}^*(h)$ is the set of points (s,y) for which*

$$\int_s^T \int_{\mathcal{O}} \frac{1}{(t-s)^{N/2}} \exp\left(-\frac{C|x-y|^2}{t-s}\right)|h(t,x,0,0)|^2 \,\mathrm{d}x\,\mathrm{d}t$$

is infinite. This set $\mathcal{N}^*(h)$ does not depend on T.

Proof. This proof is immediate using the upper bound of the Aronson estimate (11). The fact that $\mathcal{N}^*(h)$ does not depend on T is immediate, since for any $\delta > 0$, $\int_{s+\delta}^T \int_{\mathcal{O}} (t-s)^{-N/2}\exp\left(-\frac{C|x-y|^2}{t-s}\right)|h(t,x,0,0)|^2\,\mathrm{d}x\,\mathrm{d}t \leq \delta^{-N/2}\|h(\cdot,\cdot,0,0)\|_{\mathrm{L}^{2,2}(s,T;\mathcal{O})}$. □

With the martingale representation theorem (theorem 2), one knows that if $(s,y)\notin\mathcal{N}(h)$, then there exists a unique solution (Y,Z) to the BSDE (see for example Pardoux, 1999)

$$\begin{cases} Y_t = g(X_T)\mathbf{1}_{\{T\leq\tau\}} + \displaystyle\int_{t\wedge\tau}^{T\wedge\tau} h(r,X_r,Y_r,Z_r)\,\mathrm{d}r \\ \qquad\qquad\qquad - \displaystyle\int_{t\wedge\tau}^{T\wedge\tau} Z_r\,\mathrm{d}M_r,\ t\in[s,T],\ \mathbb{P}_{s,y}, \\ \mathbb{E}_{s,y}\left[\displaystyle\sup_{t\in[s,T\wedge\tau]}|Y_t|^2 + \int_s^{T\wedge\tau}\|Z_t\|^2\,\mathrm{d}t\right] < +\infty, \\ Y \text{ and } Z \text{ are } \mathcal{F}_{s,\cdot}\text{-progressively measurable,} \end{cases} \tag{32}$$

where τ is the first exit time from $\mathcal{O}$ of the process X.

If $h(\cdot,\cdot,0,0)$ belongs to $\mathrm{L}^{2,2}(0,T;\mathcal{O})$, then u is not necessarily continuous. Yet if $(Y^{s,y},Z^{s,y})$ is the unique solution of the BSDE (32) under $\mathbb{P}_{s,y}$ for $(s,y)\notin\mathcal{N}(h)$, then we set $\widehat{u}(s,y)=Y_s^{s,y}$. If $h(\cdot,\cdot,0,0)$ belongs to $\mathrm{L}^{q,p}(0,T;\mathcal{O})$, where (q,p) satisfies (10a), then $\widehat{u}$ denotes the continuous version of u.

Proposition 5. *Let u be the solution in $\mathcal{W}_{0,T}$ of the semi-linear PDE* (30), *and $\widehat{u}$ be constructed as previously. Then $\widehat{u}$ is a version of u. Moreover, for any $(s,y)\notin\mathcal{N}(h)$, $\mathbb{P}_{s,y}$, if (Y,Z) is the solution of the BSDE* (32)*, then*

$$\text{for any } t\in[s,T],\ Y_t=\widehat{u}(t,X_t) \text{ and } Z_t=\nabla u(t,X_t) \tag{33}$$

with the convention that $\widehat{u}(t,X_t)=0$ and $\nabla u(t,X_t)=0$ when $t\geq\tau$.

In fact, the identification of Z_t with $\nabla u(t,X_t)$ is not clear because $\frac{\partial u}{\partial x_i}$ belongs only to $\mathrm{L}^{2,2}(s,T;\mathbb{R}^N)$. But what is really proved is that $\int_s^T\|\nabla u(r,X_r)-Z_r\|^2\,\mathrm{d}r=0$. Then, the martingales $\int_s^\cdot\nabla u(r,X_r)\,\mathrm{d}M_r$ and $\int_s^\cdot Z_r\,\mathrm{d}M_r$ are indistinguishable on $[s,T]$. Moreover, $\int_s^\cdot h(r,X_r,u(r,X_r),\nabla u(r,X_r))\,\mathrm{d}r$ and $\int_s^\cdot h(r,X_r,Y_r,Z_r)\,\mathrm{d}r$ are also indistinguishable on $[s,T]$.

If $h(\cdot,\cdot,0,0)$ belongs to $\mathrm{L}^{2,2}(0,T;\mathcal{O})$ but not to $\mathrm{L}^{q,p}(0,T;\mathbb{R}^N)$, it is immediate from (33) that there exists a version $\widehat{u}$ of the solution u of (30) such that $t\in[s,\tau)\mapsto\widehat{u}(t,X_t)$ is continuous $\mathbb{P}_{s,y}$-almost surely for $(s,y)\notin\mathcal{N}(h)$, since $Y_t=\widehat{u}(t,X_t)$ and $t\mapsto Y_t$ is continuous.

Remark 3. The previous Proposition proves also the Itô formula for time-inhomogeneous processes generated by divergence form operators: If $u\in\mathcal{W}_{0,T}$ is such that $f=\frac{\partial u}{\partial t}+Lu$ belongs to $\mathrm{L}^{2,2}(0,T;\mathcal{O})$, then for any $(s,y)\notin\mathcal{N}(f)$, $\mathbb{P}_{s,y}$-almost surely.

$$u(t\wedge\tau,X_{t\wedge\tau}) = u(s,y) + \int_s^{t\wedge\tau} f(r,X_r)\,\mathrm{d}r + \int_s^{t\wedge\tau}\nabla u(r,X_r)\,\mathrm{d}M_r, \tag{34}$$

for any $t\in[s,T]$.

Remark 4. The problem of starting points for which the BSDE (32) may be solved is also discussed in Bally et al. (2005) in a more general setting for a time-homogeneous process. Yet, our approach is more elementary but relies on the same underlying idea: the Itô formula under a distribution $\mathbb{Q}$ is equivalent in some sense to a representation theorem under $\mathbb{Q}$.

Proof. We may assume without loss of generality that $\mathcal{O} = \mathbb{R}^N$. We assume in a first time that g and $h(\cdot,\cdot,0,0)$ are bounded. Let $(a^n, b^n)_{n\in\mathbb{N}}$ be a sequence of smooth approximations of (a, b), and $(h^n)_{n\in\mathbb{N}}$ be a sequence of smooth approximations of $h(t, x, u(t, x), \nabla u(t, x))$ which belongs to $\mathrm{L}^{2,2}(0, T; \mathbb{R}^N)$. Let u^n be the solution of the linear PDE $\frac{\partial u^n(t,x)}{\partial t} + L^n u^n(t, x) = -h^n(t, x)$ with the final condition $u^n(T, x) = g(x)$.

The proof given in Lejay, 2002a, which relies on Proposition 1 is easily extended to time-inhomogeneous operator with a differential first-order term. Hence, for any measure ν with a bounded density, one obtains that $\mathbb{P}_{s,\nu}$-almost surely, for any $t \in [s, T]$,

$$u(t, X_t) = g(X_T) + \int_t^T h(r, X_r, u(r, X_r), \nabla u(r, X_r))\,\mathrm{d}r - \int_t^T \nabla u(r, X_r)\,\mathrm{d}M_r. \tag{35}$$

Let (s, y) belongs to $[0, T) \times \mathbb{R}^N$. For any δ such that $s < s + \delta < T$, the Markov properties implies that (35) defined under $\mathbb{P}_{s+\delta,\nu}$ with $\nu(\mathrm{d}x) = \Gamma(s, y, s + \delta, x)\,\mathrm{d}x$ is also valid under $\mathbb{P}_{s,y}$, but only for $t \in [s + \delta, T]$. The boundedness of $h(t, x, 0, 0)$ and g implies that u is bounded. So, $\mathbb{E}_{s,\nu}\left[\sup_{t\in[s,T]} |u(t, X_t)|^2\right]$ is finite and then $\mathbb{E}_{s,y}\left[\sup_{t\in[s+\delta,T]} |u(t, X_t)|^2\right]$ is finite. Also, $\mathbb{E}_{s,y}\left[\int_{s+\delta}^T \|\nabla u(t, X_t)\|^2\,\mathrm{d}t\right]$ is finite.

However, using the martingale representation theorem 2, one knows that there exists a unique solution $(Y, Z)_{t\in[s,t]}$ to the BSDE (32) under $\mathbb{P}_{s,y}$. Hence, using the continuity of u and that of Y, $u(t, X_t) = Y_t$ for any $t \in [s + \delta, T]$. Besides, $\mathbb{E}_{s,y}\left[\int_{s+\delta}^T \|\nabla u(t, X_t) - Z_t\|^2\,\mathrm{d}t\right] = 0$. As u is continuous, $u(s + \delta, X_{s+\delta})$ converges to $u(s, y)$ as δ decreases to 0. Moreover, $\mathbb{E}_{s,y}\left[\int_s^T \|Z_t\|^2\,\mathrm{d}t\right]$ is finite. So, $\int_{s+\delta}^T \nabla u(r, X_r)\,\mathrm{d}M_r$ converges almost surely to $\int_s^T Z_r\,\mathrm{d}M_r = \int_s^T \nabla u(r, X_r)\,\mathrm{d}M_r$ as δ decreases to 0, and $\int_s^\cdot \nabla u(r, X_r)\,\mathrm{d}M_r$ is a square-integrable martingale under $\mathbb{P}_{s,y}$. Similarly, as δ decreases to 0, $\int_{s+\delta}^T h(r, X_r, u(r, X_r), \nabla u(r, X_r))\,\mathrm{d}r$ converges to $\int_s^T h(r, X_r, u(r, X_r), \nabla u(r, X_r))\,\mathrm{d}r$ and (35) is valid under $\mathbb{P}_{s,y}$ for any $t \in [s, T]$.

For a term $h(\cdot,\cdot,0,0)$ in $\mathrm{L}^{2,2}(0, T; \mathbb{R}^N)$, one has simply to combine this argument with the one given in Lejay, 2002a: Let (s, y) be a point in $[0, T] \times \mathbb{R}^N$ such that there exists a unique solution (Y, Z) to the BSDE (32).

We set $h(t, x) = h(t, x, u(t, x), \nabla u(t, x))$ and $h^n(t, x) = (-n) \vee h(t, x) \wedge n$. Let (Y^n, Z^n) be the solution of the BSDE (32), where h is replaced by h^n.

It follows from the Gronwall Lemma that for any $t \in [s,T]$, Y_t^n converges in $\mathrm{L}^2(\mathbb{P}_{s,y})$ to Y_t and $\mathbb{E}_{s,y}\left[\int_t^T \|Z_r^n - Z_r\|^2 \,\mathrm{d}r\right]$ decreases to 0 if

$$\int_{\mathbb{R}^N}\int_s^T \Gamma(s,y,t,x)|h^n - h|^2(t,x)\,\mathrm{d}t\,\mathrm{d}x = \mathbb{E}_{s,y}\left[\int_s^T |h^n - h|^2(t,X_t)\,\mathrm{d}t\right] \xrightarrow[n\to\infty]{} 0. \tag{36}$$

But $\int_{\mathbb{R}^N}\int_s^T \Gamma(s,y,t,x)|h(t,x)|^2\,\mathrm{d}t\,\mathrm{d}x$ is finite, and $h^n(t,x)$, which is bounded by $|h(t,x)|$, converges almost surely to $h^n(t,x)$. By the Lebesgue Dominated Convergence Theorem, (36) holds. Besides, h^n also converges to h in $\mathrm{L}^{2,2}(s,T;\mathbb{R}^N)$, and, if u^n is the solution of the PDE $\frac{\partial u^n}{\partial t} + Lu^n + h^n = 0$ with the final condition g, then $u^n(t,X_t)$ converges to $u(t,X_t)$ in $\mathrm{L}^2(\mathbb{P}_{s,y})$ for any $t \in [s,T]$, while $\int_{s+\delta}^T \|\nabla u^n - \nabla u\|^2(r,X_r)\,\mathrm{d}r$ converges in $\mathrm{L}^2(\mathbb{P}_{s,y})$ to 0 for any $\delta > 0$. As previously, Z_t may be identified with $\nabla u(t,X_t)$. Concerning Y_t, one remarks that for any $t > s$, $u(t,X_t) = Y_t$ $\mathbb{P}_{s,y}$-almost surely.

Since $u^n(s,x)$ converges in $\mathrm{L}^2(\mathbb{R}^N)$ to $u(s,x)$, a subsequence of $(u^n(s,\cdot))_{n\in\mathbb{N}}$ converges almost everywhere to $u(s,\cdot)$. On the other hand, since u^n is continuous, $u^n(s,y) = Y_s^{n,s,y}$, where $(Y^{n,s,y}, Z^{n,s,y})$ is the solution of (32) under $\mathbb{P}_{s,y}$ with h replaced by h^n. From standard computations, $Y^{n,s,y}$ converges to $Y^{s,y}$, where $(Y^{s,y}, Z^{s,y})$ is the solution of (32) under $\mathbb{P}_{s,y}$. So, $y \mapsto Y_y^{s,y} = \widehat{u}(s,y)$ is a version of $u(s,\cdot)$. Finally, it is standard that $\mathbb{P}_{s,y}$-almost surely, for any $t \in [s,T]$, $Y_t^{s,y} = Y_t^{t,X_t^{s,y}}$, so that $Y_t^{s,y} = \widehat{u}(t,X_t^{s,y})$, where $X^{s,y}$ denotes the process X under $\mathbb{P}_{s,y}$.

For a general final condition $g \in \mathrm{L}^2(\mathbb{R}^N)$, the proof is similar. □

2.2 Quasi-linear PDEs

For a function u in $\mathcal{W}_{0,T}$, we define by L^u the divergence form operator by

$$L^u = \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{i,j}(t,x,u(t,x))\frac{\partial}{\partial x_j}\right) + b_i(t,x,u(t,x),\nabla u(t,x))\frac{\partial}{\partial x_i},$$

where a and b satisfies: For any $(t,x,y,z) \in [0,T]\times\mathcal{O}\times\mathbb{R}\times\mathbb{R}^N$,

- ⋆ $a(t,x,y)$ is a $N\times N$-symmetric matrix and $b(t,x,y,z) \in \mathbb{R}^N$, (37a)
- ⋆ $(t',x',y',z') \mapsto \psi(t',x',y',z')$ is measurable for $\psi = a, b$, (37b)
- ⋆ $\lambda\|z\|^2 \leq a(t,x,y)z\cdot z$, (37c)
- ⋆ $a_{i,j}(t,x,y) \leq \Lambda$, (37d)
- ⋆ $(y',z') \mapsto \psi(t,x,y',z')$ is continuous with $\psi = a, b$, (37e)
- ⋆ $\|b(t,x,y,z)\| \leq \Lambda$. (37f)

We are now interested by solving in $\mathcal{W}_{0,T}$,

$$\frac{\partial u}{\partial t}(t,x) + L^u u(t,x) + h(t,x,u(t,x),\nabla u(t,x)) = 0 \tag{38}$$

with the final condition $u(T,x) = g(x)$. The conditions on h and g are still (31a)–(31e).

Theorem 3. *We assume that $\mathcal{O}$ is bounded. Under the Hypotheses* (37a)–(37f) *and* (31a)–(31e)*, there exists a weak solution to the parabolic PDE* (38)*.*

This Theorem is proved in Ladyženskaja et al., 1968, Theorem V.6.7, p. 466, using the Leray-Schauder fixed point theorem. The Lipschitz assumption on $(y,z) \mapsto h(t,x,y,z)$ is too strong here, but is used when dealing with BSDEs. This Theorem does not provide uniqueness of the solution, unless one assumes that the coefficients are more regular.

If u belongs to $\mathcal{W}_{0,T}$, it is clear that the coefficients of L^u satisfies (7a)–(7e). So, this differential operator L^u generates a continuous semi-group $(P^u_{s,t})_{t\geq s}$ and a strong Markov process X^u.

Theorem 4. (i) *Let u be a weak solution in $\mathcal{W}_{0,T}$ of the quasi-linear PDE* (38)*. Then, for any $(s,y) \notin \mathcal{N}(h)$, the unique solution (Y^u_t, Z^u_t) of the BSDE* (32) *where X is replaced by X^u and M by the martingale part M^u of X^u, is equal to $(u(t,X_t), \nabla u(t,X_t))$ on $[s,T]$. Moreover, u is also a* mild solution *of* (38)*, i.e.,*

$$u(s,x) = P^u_{s,T}g(x) + \int_s^T P^u_{s,r}h(r,x,u(r,x),\nabla u(r,x))\,\mathrm{d}r \tag{39}$$

for any $(s,x) \in \mathcal{N}(h)$.

(ii) *Let v be a function in $\mathcal{W}_{0,T}$ such that the process X^v satisfies for any $(s,y) \notin \mathcal{N}(h)$, $\mathbb{P}_{s,y}$-almost surely,*

$$\begin{aligned} v(t\wedge\tau, X^v_{t\wedge\tau}) = g(X^v_T)\mathbf{1}_{\{T\leq\tau\}} + \int_{t\wedge\tau}^{T\wedge\tau} h(r,X^v_r,v(r,X^v_r),\nabla v(r,X^v_r))\,\mathrm{d}r \\ - \int_{t\wedge\tau}^{T\wedge\tau} \nabla v(r,X^v_r)\,\mathrm{d}M^v_r,\ t\in[s,T], \end{aligned} \tag{40}$$

where M^v is the martingale part of X^v. Then, v is a weak solution of (38)*.*

The proof of (ii) relies on the following result.

Proposition 6. *A mild solution of* (38)*, i.e., a function $u \in \mathcal{W}_{0,T}$ satisfying* (39)*, is also a weak solution of* (38)*.*

Proof. Let $(P^u_{s,t})_{t\geq s}$ be the semi-group generated by L^u, for some $u\in\mathcal{W}_{0,T}$. If w belongs to $\mathrm{H}^1_0(\mathcal{O})$ and if $f\in\mathrm{L}^{2,2}(0,T;\mathcal{O})$ is continuous on $[0,T]\times\mathcal{O}$, then the function $s\mapsto\langle P^u_{s,t}f(t,\cdot),w\rangle$ is differentiable on $(0,t)$ for almost every s, and its derivative is $\mathcal{L}_s(P^u_{s,t}f(t,\cdot),w)$, where $\mathcal{L}_s$ is the bilinear form on $\mathrm{H}^1_0(\mathcal{O})\times\mathrm{H}^1_0(\mathcal{O})$ defined by

$$\mathcal{L}_s(v,w)=\frac{1}{2}\int_{\mathcal{O}}a(s,x,u(s,x))\nabla v(x)\cdot\nabla w(x)\,\mathrm{d}x \\ -\int_{\mathcal{O}}b(s,x,u(s,x),\nabla u(s,x))\cdot\nabla v(x)\,w(x)\,\mathrm{d}x.$$

We set $u(s,x)=P^u_{s,T}g(x)+\int_s^T P^u_{s,t}f(t,x)\,\mathrm{d}t$.

Let w be a function in $\mathrm{H}^1_0(\mathcal{O})$, and s be a real in $(0,T)$. Hence,

$$\begin{aligned}\frac{\partial}{\partial s}\langle u(s,\cdot),w\rangle&=\frac{\partial}{\partial s}\left\langle P^u_{s,T}g,w\right\rangle+\frac{\partial}{\partial s}\int_s^T\left\langle P^u_{s,t}f(t,\cdot),w\right\rangle\,\mathrm{d}t\\&=\mathcal{L}_s(P^u_{s,T}g,w)+\int_s^T\mathcal{L}_s(P^u_{s,t}f(t,\cdot),w)\,\mathrm{d}t-\langle f(s,\cdot),w\rangle\\&=\mathcal{L}_s(u(s,\cdot),w)-\langle f(s,\cdot),w\rangle\,.\end{aligned}$$

By density, this is true for any function f in $\mathrm{L}^{2,2}(0,T;\mathcal{O})$, and we may use this result for $f(t,x)=h(t,x,u(t,x),\nabla u(t,x))$. This is true for almost every $s\in[0,T)$, and then u is weak solution to (38). □

Proof of Theorem 4. The first assertion is a direct consequence of Proposition 5. If $(s,y)\notin\mathcal{N}(h)$, then $\mathbb{E}_{s,y}\left[\int_s^{T\wedge\tau}\|\nabla u\|^2(r,X^u_r)\,\mathrm{d}r\right]$ is finite and $\int_s^{\cdot\wedge\tau}\nabla u(r,X^u_r)\,\mathrm{d}r$ is a martingale. As Y^u_s is equal to $u(s,y)$. So,

$$u(s,y)=\mathbb{E}_{s,y}\left[\,g(X^u_T);T\leq\tau\,\right]+\mathbb{E}_{s,y}\left[\int_s^{T\wedge\tau}h(r,X^u_r,u(r,X^u_r),\nabla u(r,X^u_r))\,\mathrm{d}r\right].$$

The Fubini theorem applied to the last equation leads to (39), since by definition $\mathbb{P}_{s,y}\left[\,f(X^u_t);t<\tau\,\right]=P^u_{s,t}f(y)$.

The proof of (ii) is immediate since applying the expectation $\mathbb{E}_{s,y}$ on each side of (40) implies that v is a mild solution and then, according to Proposition 6, a weak solution of (38). □

3 Weak solution of FBSDE

We have seen, and this is not surprising, that it is possible to associate a BSDE to a quasi-linear PDE. However, the weakness of this representation with respect to the representation by FBSDEs is that the process X itself depends on the choice of a solution u. With Theorem 1, it is possible to assert that (X, Y, Z) is the solution of some FBSDE, but this equation involves the transition density function Γ of the process itself and two Brownian motions, one being adapted to $\mathcal{F}$, but the other to $\mathcal{F}^{\overline{X}}$. Thus, we are now interested in processes generated by divergence form operators, but which are also solutions of some SDE. For that, the diffusion coefficient shall be differentiable.

We will first see what happens when we transform a quasi-linear PDE with a divergence form to a PDE with a non-divergence form operator, and we introduce the notion of weak solution of FBSDEs.

In a second part, we proceed according to an inverse method: we start first from a quasi-linear PDE with a non-divergence form operator, and transform it to the solution of a quasi-linear PDE. A quadratic term in the gradient of the solution appears. However, there are some cases where the solution u remains in $\mathcal{W}_{0,T}$, so that our previous results may be applied.

3.1 Weak solution of FSBDE

If the coefficient a is smooth enough, and if u is differentiable, then

$$\frac{\partial a_{i,j}(t,x,u(t,x))}{\partial x_i} = \left(\frac{\partial a_{i,j}}{\partial x_i} + \frac{\partial a_{i,j}}{\partial y}\frac{\partial u}{\partial x_i}\right)(t,x,u(t,x)).$$

So, one may transform a divergence form operator into a non-divergence form operator.

In addition to (31a)–(31e) and (37a)–(37f), we assume that:

$$\star\ a_{i,j}(t,\cdot,\cdot) \in \mathrm{W}^{1,\infty}(\mathcal{O}\times\mathbb{R}),\ \forall t\in[0,T), \tag{41a}$$

$$\star\ y \mapsto a(t,x,y) \in \mathcal{C}^1(\mathbb{R}) \text{ for any } (t,x)\in[0,T)\times\mathcal{O}, \tag{41b}$$

$$\star\ h(t,x,0,0) \text{ is bounded}, \tag{41c}$$

$$\star\ \operatorname{ess\,sup}_{(t,x,y)\in[0,T]\times\mathcal{O}\times\mathbb{R}} \left|\left(\frac{\partial a_{i,j}}{\partial x_i}, \frac{\partial a_{i,j}}{\partial y}\right)\right|(t,x,y) < +\infty, \tag{41d}$$

$$\star\ g \in \mathrm{W}^{1,\infty}(\mathcal{O}) \cap \mathrm{H}^1_0(\mathcal{O}). \tag{41e}$$

Remark 5. The previous hypotheses on the coefficient a imply that it is Lipschitz in x, y, uniformly in each of its variables.

Remark 6. The assumption that g is weakly differentiable and in $\mathrm{W}^{1,\infty}(\mathcal{O})$ is to avoid the explosion of $\nabla u(t,x)$ as t increases to T.

Proposition 7. *Under the hypotheses* (31a)–(31e)*,* (37a)–(37f) *and* (41a)–(41e)*, the gradient ∇u of the solution u of* (38) *is bounded on the cylinder $[0,T]\times\mathcal{O}$. Furthermore, $\frac{\partial u}{\partial t}$ and $\frac{\partial^2 u}{\partial x_i \partial x_j}$ belong to $\mathrm{L}^{2,2}(0,T;\mathcal{O})$.*

Proof. Let $\mathcal{O}$ be a bounded open set with a smooth boundary. Let X^u be the process generated by L^u. With (41c), $\mathcal{N}(h) = \emptyset$, and then the solution (Y^u, Z^u) of the BSDE (32) with respect to X^u is defined for any $(s, x) \in [0, T) \times \mathcal{O}$. For a given point $(s, x) \in [0, T) \times \mathcal{O}$, standard computations on BSDEs (See for example Proposition 2.2 and Theorem 4.1 in Pardoux, 1999, p. 509 and p. 526) implies that $u(s, x) = Y^u_s$ is bounded by some constant that depends only on T and on the bounds of $(t, x) \mapsto h(t, x, 0, 0)$ and g. So, u is globally bounded on $\mathcal{O}$.

Theorem V.4.1 in Ladyženskaja et al., 1968, p. 443, proves the boundedness of ∇u, and the fact that $\frac{\partial u}{\partial t}$ and $\frac{\partial^2 u}{\partial x_i \partial x_j}$ are in $\mathrm{L}^{2,2}(0, T; \mathcal{O})$. But for simplicity, this theorem is proved for classical solutions in $\mathcal{C}^{1,2}([0, T] \times \mathcal{O})$. However, there is no difficulty to prove the result for generalized solutions. Once u is given, we freeze the coefficients of the quasi-linear PDE as previously, so that u is the solution of a linear PDE

$$\frac{\partial u}{\partial t}(t, x) + L^u u(t, x) + f(t, x) = 0,$$

with $f(t, x) = h(t, x, u(t, x), \nabla u(t, x)) \in \mathrm{L}^{2,2}(0, T; \mathcal{O})$. We consider now a regularization of the coefficients $a(t, x, u(t, x))$ and $b(t, x, u(t, x), \nabla u(t, x))$ of L^u, and a regularization of f. Then, the solution u^n to the corresponding PDE is a classical solution. By Theorem V.3.1 in Ladyženskaja et al., 1968, p. 437, the upper bound of ∇u^n is bounded by some value C depending only on the constants that appear in the hypotheses (31a)–(31e), (37a)–(37f) and (41a)–(41e), some geometric properties of the boundary of $\mathcal{O}$, and the upper bound of u (which itself depends only g and h). So, $|\nabla u^n(t, x)| \leq C$ for any $n > 0$ and any $(t, x) \in [0, T] \times \mathcal{O}$. But ∇u^n converges to ∇u in $\mathrm{L}^{2,2}(0, T; \mathcal{O})$. Hence $|\nabla u(t, x)| \leq C$ for any $(t, x) \in [0, T] \times \mathcal{O}$.

We assume now that $\mathcal{O} = \mathbb{R}^N$, Let $B(y, R)$ be the ball of radius R and centered on y, and $\mathcal{Q}^{y,R}_T = [0, T] \times B(y, R)$. In this case, $\|\partial u/\partial t\|_{\mathrm{L}^2(\mathcal{Q}^{y,R}_T)}$ and $\sup_{(t,x)\in\mathcal{Q}^{y,R}_T} \|\nabla u(t, x)\|$ depend only on the bound of u on $\mathcal{Q}^{y,2R}_T$, the radius R, the constants of the hypotheses, and the bounds of g and its derivative on $\mathcal{Q}^{y,2R}_T$. However, by (41e), both g and ∇g are globally bounded, and we have seen that u is globally bounded. So, it is now clear that ∇u is also globally bounded, since its bound on $\mathcal{Q}^{y,R}_T$ does not depend on y.

Similarly, the norms in $\mathrm{L}^{2,2}(0, T; \mathcal{O})$ of $\frac{\partial u^n}{\partial t}$ and $\frac{\partial^2 u^n}{\partial x_i \partial x_j}$ are also bounded by some values that depend only on the constants of the hypotheses and the geometry of $\mathcal{O}$, and so are their limits. □

Remark 7. The solution u is continuous, and its derivative ∇u is itself Hölder continuous on every subdomain of the cylinder $[0, T] \times \mathcal{O}$ separated from its boundary by a positive distance. But u is in general not of class $\mathcal{C}^2$.

It is now time to give a definition of weak solutions, which is similar to the notion of weak solution of SDEs. Definition 1 below concerns FBSDE like the following one:

$$\begin{cases} X_{t\wedge\tau} = x + \int_s^{t\wedge\tau} \sigma(r, X_r, Y_r)\,\mathrm{d}B_r + \int_s^{t\wedge\tau} \widehat{b}(r, X_r, Y_r, Z_r)\,\mathrm{d}r, \\ Y_t = g(X_T)\mathbf{1}_{\{T\leq\tau\}} + \int_{t\wedge\tau}^{T\wedge\tau} h(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_{t\wedge\tau}^{T\wedge\tau} Z_r\sigma(r, X_r, Y_r)\,\mathrm{d}B_r, \end{cases} \tag{42}$$

where B is a Brownian Motion and σ satisfies

$$\star\ \sigma(t,x,y) \text{ is a } m\times N\text{-matrix for any } (t,x,y)\in[0,T]\times\mathcal{O}\times\mathbb{R}, \tag{43a}$$

$$\star\ \sigma(t,x,y)\sigma(t,x,y)^{\mathrm{T}} = a(t,x,y) \text{ for any } (t,x,y)\in[0,T]\times\mathcal{O}\times\mathbb{R}, \tag{43b}$$

$$\star\ \sigma \text{ is measurable.} \tag{43c}$$

Definition 1. A *weak solution* of the FBSDE (42) is a family of probabilities $(\Omega, \mathcal{G}_\infty, \mathbb{P}_{s,y}, s\in[0,T), y\in\mathcal{O})$ with a filtration $(\mathcal{G}_{s,\cdot})_{s\in[0,T]} = (\mathcal{G}_{s,t})_{0\leq s\leq t\leq T}$ and some stochastic processes B, X, Y, Z such that under $\mathbb{P}_{s,y}$, for any $(s,y)\in[0,T)\times\mathcal{O}$,

$$\star\ B, X, Y, Z \text{ are } \mathcal{G}_{s,\cdot}\text{-progressively measurable,} \tag{44a}$$

$$\star\ B \text{ is a } m\text{-dimensional } \mathcal{G}_{s,\cdot}\text{-Brownian motion,} \tag{44b}$$

$$\star\ X_t, Y_t, Z_t \text{ take their values in } \mathbb{R}^N,\ \mathbb{R} \text{ and } \mathbb{R}^N, \tag{44c}$$

$$\star\ \tau = \inf\{\, t\geq s \,|\, X_t\notin\mathcal{O}\,\} \text{ is a } \mathcal{G}_{s,\cdot}\text{-stopping time,} \tag{44d}$$

$$\star\ \mathbb{E}_{s,y}\left[\sup_{s\leq t\leq T\wedge\tau}\left(|X_t|^2+|Y_t|^2\right) + \int_s^{T\wedge\tau}\|Z_t\|^2\,\mathrm{d}t\right] < +\infty, \tag{44e}$$

$$\star\ \mathbb{P}_{s,y}\text{-almost surely, (42) holds for any } t\in[s,T]. \tag{44f}$$

Proposition 8. *Let us consider σ satisfying* (43a)–(43c)*, and such that $a = \sigma\sigma^{\mathrm{T}}$, b, h and g satisfy Hypotheses* (31a)–(31e), (37a)–(37f) *and* (41a)–(41e)*. Then there exists a weak solution to the FBSDE* (42) *with*

$$\widehat{b}_i(t,x,y,z) = \left[\frac{1}{2}\frac{\partial a_{k,i}}{\partial x_k} + \frac{1}{2}\frac{\partial a_{k,i}}{\partial y}z_k + b_i\right](t,x,y,z).$$

Proof. Let u be a solution of (38). The operator L^u generates a stochastic process $(\Omega, \mathcal{F}_\infty, \mathcal{F}_{s,t}, X_t, \mathbb{P}_{s,y}; 0\leq s\leq t\leq T, y\in\mathbb{R}^N)$ (We have to remember that the coefficients of L^u are extended to $\mathbb{R}^N$, so that X is conservative).

For any function $f\in\mathcal{C}_{\mathrm{b}}^{1,2}([0,T]\times\mathcal{O})$, $L^u f(t,x) = \frac{1}{2}a_{i,j}(t,x,u(t,x))\frac{\partial^2 f(t,x)}{\partial x_i\partial x_j} + \widehat{b}_i(t,x,u(t,x),\nabla u(t,x))\frac{\partial f(t,x)}{\partial x_i}$ on $[0,T]\times\mathcal{O}$. As ∇u is bounded, it is clear that $\widehat{b}^u(t,x) = \widehat{b}(t,x,u(t,x),\nabla u(t,x))$ is bounded. So, $\mathcal{N}(L^u f) = \emptyset$ since $L^u f$ is bounded. It follows from the Itô formula (34) applied to $f(x) = x_i$ that for any $(s,y) \in [0,T] \times \mathbb{R}^N$, $X_{t\wedge\tau} = y + M_{t\wedge\tau} + \int_s^{t\wedge\tau} \widehat{b}^u(r, X_r)\,\mathrm{d}r$, where M_t is the martingale with cross-variations $\langle M^i, M^j\rangle_t = \int_s^t a^u_{i,j}(r, X_r)\,\mathrm{d}r$. Is is also clear that $\mathbb{E}_{s,y}\left[\sup_{0\leq t\leq T\wedge\tau} |X_t|^2\right]$ is finite.

So, there exists an extension $(\widetilde{\Omega}, \mathcal{G}_\infty, \widetilde{\mathbb{P}}_{s,y})$ of the space $(\Omega, \mathcal{F}_\infty, \mathbb{P}_{s,y})$, a filtration $\mathcal{G}_{s,\cdot}$ and a $(\widetilde{\mathbb{P}}_{s,y}, \mathcal{G}_{s,t}; t \geq s)$-Brownian Motion B (See for example Proposition 5.4.6 in Karatzas and Shreve, 1991, p. 315) such that $M_t = \int_s^t \sigma(r, X_r, u(r, X_r))\,\mathrm{d}B_r$, $\widetilde{\mathbb{P}}_{s,y}$-almost surely. Moreover, M and X are also $\mathcal{G}_{s,\cdot}$-adapted, and τ is an $\mathcal{G}_s$-stopping time (see Remark 3.4.1 in Karatzas and Shreve, 1991, p. 169).

We set $Y_t = u(t, X_t)$ and $Z_t = \nabla u(t, X_t)$. With (41c), (Y, Z) is solution to the BSDE (32) under $\mathbb{P}_{s,y}$ for any $(s, y) \in [0,T] \times \mathcal{O}$. So, the result is proved by substituting Y_t and Z_t to $u(t, X_t)$ and $\nabla u(t, X_t)$ both in the expression of X and in the BSDE that (Y, Z) solves. □

Remark 8. No continuity in (t, x) is required for the first-order differential term b nor the non-linear term h. This is why this result is different from the results on strong solutions of FBSDEs, where Lipschitz assumptions in x on σ, b and h are crucial.

Remark 9. Even if it is known that the solution u of the quasi-linear PDE (38) is unique, nothing proves us that the solution (B, X, Y, Z) of (42) is also unique. Generally, when (X, Y, Z) is the solution of an FBSDE, it is an open problem to know whether or not there exists a deterministic, measurable function v on $[0,T]\times\mathcal{O}$ such that $Y_t = v(t\wedge\tau, X_{t\wedge\tau})$ and $Z_t = \nabla v(t\wedge\tau, X_{t\wedge\tau})$.

Remark 10. The previous results may be extended to the case where a is also non-linear in ∇u. However, when one transforms the divergence form operator into a non-divergence form operator, then one sees a term $\frac{\partial^2 u}{\partial x_i \partial x_j}$ appearing. These terms generally belong to $\mathrm{L}^{2,2}(0,T;\mathcal{O})$. Hence, we are led to introduce supplementary terms of type $U_t^{i,j} = \frac{\partial^2 u}{\partial x_i \partial x_j}(t, X_t)$ for $i, j = 1, \dots, N$. Besides, $t \mapsto \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x)$ is not necessarily bounded as t goes to T, so $U_t^{i,j}$ is not necessarily defined under $\mathbb{P}_{s,y}$ for any (s, y).

3.2 Weak solution of FBSDE for a non-divergence form operator

The drift term $\widehat{b}$ in Proposition 8 does not give us an entire satisfaction, since it involves the term a. One might ask if, given σ, b and h, it is possible to assert the existence of a weak solution to the FBSDE (42) using our previous scheme. Under the condition of existence of weak solutions for quasi-linear

PDE with a non-linear term in ∇u having a quadratic growth, the answer to this question is positive.

Proposition 9. *Let us consider σ satisfying* (43a)–(43c) *and such that $a = \sigma\sigma^{\mathrm{T}}$, b, h and g satisfy Hypotheses* (31a)–(31e), (37a)–(37f) *and* (41a)–(41e).

We assume that there exists a solution u in $\mathcal{W}_{0,T} \cap \mathrm{L}^\infty([0,T) \times \mathcal{O})$ to the quasi-linear PDE

$$\frac{\partial u}{\partial t}(t,x) + \widehat{L}^u u(t,x) + h(t,x,u(t,x),\nabla u(t,x)) = 0 \text{ and } u(T,x) = g(x) \quad (45)$$

with

$$\widehat{L}^u = \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{i,j}(t,x,u(t,x))\frac{\partial}{\partial x_j}\right) - \frac{1}{2}\frac{\partial a_{i,j}}{\partial x_j}(t,x,u(t,x))\frac{\partial}{\partial x_i} \\ - \frac{1}{2}\frac{\partial a_{i,j}}{\partial y}(t,x,u(t,x))\frac{\partial u}{\partial x_j}(t,x)\frac{\partial}{\partial x_i} + b_i(t,x,u(t,x))\frac{\partial}{\partial x_i}.$$

Then there exists a weak solution (B,X,Y,Z) to the FBSDE (42) *with $\widehat{b} = b$.*

Proof. Under the assumption that there exists a bounded solution to (45), Theorem V.3.1 and Theorem V.4.1 in Ladyženskaja et al., 1968 may still be applied. If $\mathcal{O}$ is bounded with a smooth boundary, this means that ∇u is bounded on $[0,T]\times\mathcal{O}$. Hence, $\frac{\partial a_{i,j}}{\partial y}(t,x,u(t,x))\frac{\partial u}{\partial x_i}$ is bounded, and this means that the coefficients of $\widehat{L}^u$ satisfy (7a)–(7e). Let us denote by X the stochastic process generated by $\widehat{L}^u$. For any (s,y), $(Y_t,Z_t) = (u(t,X_t),\nabla u(t,X_t))$ is solution to (32) under $\mathbb{P}_{s,y}$. On the other hand, the infinitesimal generator of X is $\frac{1}{2}a_{i,j}(t,x,u(t,x))\frac{\partial^2}{\partial x_i\partial x_j} + b_i(t,x,u(t,x),\nabla u(t,x))\frac{\partial}{\partial x_i}$. Thus, it remains to conclude as in the proof of Proposition 8.

Let us assume now that $\mathcal{O} = \mathbb{R}^N$. Let (s,y) be fixed in $[0,T]\times\mathbb{R}^N$. For each $R > 0$, Theorem V.4.1 in Ladyženskaja et al., 1968 asserts that the gradient ∇u is bounded on each strip of type $[0,T]\times B(y,R)$. Let $(X^R,\mathbb{P}^R_{s,y})$ be the process generated by

$$\widehat{L}^u = \frac{1}{2}\frac{\partial}{\partial x_i}\left(a_{i,j}(t,x,u(t,x))\frac{\partial}{\partial x_j}\right) - \frac{1}{2}\frac{\partial a_{i,j}}{\partial x_j}(t,x,u(t,x))\frac{\partial}{\partial x_i} \\ - \mathbf{1}_{B(y,R)}(x)\frac{1}{2}\frac{\partial a_{i,j}}{\partial y}(t,x,u(t,x))\frac{\partial u}{\partial x_j}(t,x)\frac{\partial}{\partial x_i} + b_i(t,x,u(t,x))\frac{\partial}{\partial x_i}.$$

Let τ^R be the first exit time from $B(y,R)$. On $\left\{T < \tau^R\right\}$,

$$u(t,X^R_t) = g(X^R_T) + \int_t^T h(r,X^R_r,u(r,X^R_r),\nabla u(r,X^R_r))\,\mathrm{d}r - \int_t^T \nabla u(r,X^R_r)\,\mathrm{d}M^R_r,$$

where M^R is the martingale part of X^R. On the other hand $\mathbb{P}^R_{s,y}\left[T < \tau^R\right] = \mathbb{P}^R_{s,y}\left[\sup_{t\in[s,T]} |X^R_{t\wedge\tau^R} - x| \geq R\right]$. But $X^R_{t\wedge\tau^R} = y + M^R_{t\wedge\tau^R} + \int_s^{t\wedge\tau^R} b^u(r, X^R_r)\,\mathrm{d}r$ with $b^u(r,x) = b(r,x,u(r,x),\nabla u(r,x))$, which is bounded. Moreover, $\langle M^{i,R}\rangle_t \leq \Lambda t$ for $i = 1,\dots,N$. Using the boundedness of b^u and that of $\langle M^{i,R}\rangle$, it is standard that $\mathbb{P}^R_{s,y}\left[T < \tau^R\right]$ decreases to 0 as R increases to infinity. So, we can have let R increasing to infinity, which proves the result by using localization techniques. □

The hypotheses that there exists a weak solution u to (45) in the space $\mathcal{W}_{0,T} \cap \mathrm{L}^\infty([0,T)\times\mathcal{O})$ seems unnatural at first sight. Generally, a polynomial growth in the variable ∇u of order p is sought in the space of functions whose $(p+1)$-power are integrable.

On the other hand, we use PDE (45) instead of the equivalent PDE with a differential, non-divergence form operator $\frac{\partial}{\partial t} + \frac{1}{2}a_{i,j}\frac{\partial^2}{\partial x_i \partial x_j} + b_i\frac{\partial}{\partial x_i}$, because existence results for the last one generally require more regularity on the coefficients such as being continuously differentiable in all its variables, and the solution is not sought in the space $\mathcal{W}_{0,T}$.

In fact, there has been a large amount of work proving that there may exist weak solutions of (45) in $\mathcal{W}_{0,T} \cap \mathrm{L}^\infty([0,T)\times\mathcal{O})$ (see for example Boccardo et al., 1984, 1989; Orsina and Porzio, 1992; Grenon, 1993; Porretta, 1999, and references within for parabolic or elliptic cases). Some of the existence theorems rely on the existence of some sub-solutions and super-solutions, but some of them are explicit. With more stringent assumptions, uniqueness could be proved (see for example Kobylanski, 2000). The general question of uniqueness of solutions of quasi-linear PDEs is far to be solved, unless the coefficients are more regular. Any counter-example to uniqueness of solution of (45) gives also a counter-example to the uniqueness of the solutions of some FBSDE.

The general construction of solution of quasi-linear PDE with quadratic growth is in general rather complicated. We give now a simple example of conditions ensuring the existence of a weak solution to the system (45). This result relies on the transformation of the PDE to another PDE known to have a solutions, and that sort of approach is generally chosen to study solution of PDE with quadratic growth (and also to study BSDEs with quadratic growth: See Kobylanski, 2000; Gaudron, 1999, for example).

Proposition 10. *We assume that $\mathcal{O}$ is bounded and that $N = 1$. We also assume that $a(t,x,y) = a(y)$, $h(t,x,y,z) = h(t,x,y)$ and the Hypotheses* (31a)–(31e)*,* (37a)–(37f) *and* (41a)–(41e) *are satisfied. Then there exists a* *solution in* $\mathcal{W}_{0,T}$ *to*

$$\begin{aligned}\frac{\partial u}{\partial t}+\frac{1}{2}a\frac{\partial^2 u}{\partial x^2}+b(u)\frac{\partial u}{\partial x}+h(u)\\ =\frac{\partial u}{\partial t}+\frac{1}{2}\frac{\partial}{\partial x}\left(a\frac{\partial u}{\partial x}\right)-\frac{1}{2}\frac{\mathrm{d}a}{\mathrm{d}y}\left(\frac{\partial u}{\partial x}\right)^2+b(u)\frac{\partial u}{\partial x}+h(u)=0,\quad (46)\end{aligned}$$

and there exists a weak solution to the FBSDE

$$\begin{cases}X_t=x+\int_s^t\sigma(Y_r)\,\mathrm{d}B_r+\int_s^t b(r,X_r,Y_r)\,\mathrm{d}r,\\ Y_t=g(X_T)+\int_t^T h(r,X_r,Y_r)\,\mathrm{d}r-\int_t^T Z_r\sigma(Y_r)\,\mathrm{d}B_r,\end{cases}$$

where $\sigma(y)=\sqrt{a(y)}$.

Proof. Let us denote by f the inverse of the increasing function $y\mapsto\int_0^y a(y')^{-1}\,\mathrm{d}y'$. This function f satisfies

$$a(f(y))\frac{\partial^2 f}{\partial y^2}(y)=\frac{\partial a}{\partial y}(f(y))\left(\frac{\partial}{\partial y}f(y)\right)^2.$$

Hence, we remark that for an arbitrary differentiable function $u(t,x)$,

$$\begin{aligned}&\frac{\partial f(u(t,x))}{\partial t}+\frac{1}{2}a(f(u(t,x)))\frac{\partial^2 f(u(t,x))}{\partial x^2}\\ &\qquad+b(t,x,f(u(t,x)))\frac{\partial f(u(t,x))}{\partial x}+h(t,x,f(u(t,x)))\\ &=\frac{\partial f}{\partial y}(u(t,x))\frac{\partial u}{\partial t}(t,x)+\frac{\partial f}{\partial y}(u(t,x))\Bigg[\frac{1}{2}\frac{\partial}{\partial x}\left(a(f(u(t,x)))\frac{\partial u}{\partial x}\right)\\ &\quad+b(t,x,f(u(t,x)))\frac{\partial u}{\partial x}(t,x)+\left(\frac{\partial f}{\partial y}\right)^{-1}(u(t,x))h(t,x,f(u(t,x)))\Bigg].\end{aligned}$$

As f is itself Lipschitz, our hypotheses on a, h and b imply that there exists a weak solution u in $\mathcal{W}_{0,T}$ to

$$\begin{aligned}&\partial_t u(t,x)+\frac{1}{2}\partial_x(a(f(u(t,x)))\partial_x u(t,x))\\ &\qquad+b(t,x,f(u(t,x)))\partial_x u(t,x)+a(f(u(t,x)))h(t,x,f(u(t,x)))=0.\end{aligned}$$

Furthermore, with Proposition 7, u and ∇u are bounded, and $\frac{\partial u}{\partial t}$ belongs to $\mathrm{L}^{2,2}(0,T;\mathcal{O})$.

So, $f(u)$ is also in $\mathcal{W}_{0,T}$, $\nabla f(u)$ is bounded, and $\frac{\partial f(u)}{\partial t}=\frac{\partial f}{\partial y}(u)\frac{\partial u}{\partial t}$. In conclusion, $f(u)$ is solution to (46) and Proposition 9 may be applied. □

Remark 11. The results given in this article shall also be valid for quasi-linear elliptic PDEs.

Acknowledgment. The author wishes to thank the referees for their careful reading and for having suggested many improvements.

References

Antonelli, F., Ma, J., 2003. Weak solutions of Forward-Backward SDE's. *Stochastic Anal. Appl.* 31(3), 493–514. 4

Aronson, D.G., 1968. Non-negative solutions of linear parabolic equation. *Ann. Scuola Norm. Sup. Pisa* 22, 607–693. 7, 18, 21

Bally, V., Pardoux, É., Stoica, L., 2005. BSDEs associated to semilinear PDEs with measurable coefficients. *Potential Anal.* 22(1), 17–60. 3, 24

Barles, G., Lesigne, E., 1997. SDE, BSDE and PDE. In *Backward stochastic differential equations*, Pitman Research Notes in Mathematics Series, pp 47–80. Longman. 3

Blumenthal, R.M., Getoor, R.K., 1968. *Markov Process and Potential Theory.* Academic Press. 8

Boccardo, L., Murat, F., Puel, J.-P., 1984. Résultats d'existence pour certains problèmes elliptiques quasilinéaires. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) 11 (2), 213–235. 33

Boccardo, L., Murat, F., Puel, J.-P., 1989. Existence result for some quasilinear parabolic equations. *Nonlinear Anal.* 13 (4), 373–392. 33

Chen, Z.Q., Zhao, Z., 1995. Diffusion processes and second order elliptic operators with singular coefficients for lower order terms. *Math. Ann.* 302 (2), 323–357. 20

Föllmer, H., 1981. Dirichlet processes. In *Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980)*, vol. 851 of Lecture Notes in Math., pp 476–478. Springer, Berlin. 2

Fukushima, M., Oshima, Y., Takeda, M., 1994. *Dirichlet Forms and Symmetric Markov Process.* De Gruyter. 3

Gaudron, G., 1999. *Convergence en loi d'EDS et d'EDS Rétrogrades. Application à l'homogénéisation d'EDP linéaires ou semilinéaires.* PhD thesis, Université de Provence, Marseille, France. 33

Grenon, N., 1993. Existence results for some quasilinear parabolic problems. *Ann. Mat. Pura Appl.* (4) 165, 281–313. 33

Haussmann, U. G., Pardoux, É., 1986. Time reversal of diffusions. *Ann. Probab.* 14 (4), 1188–1205. 12

Karatzas, I., Shreve, S.E., 1991. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, 2nd ed. 19, 31

Kobylanski, M., 2000. Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.* 28 (2), 558–602. 33

Ladyženskaja, O.A., Solonnikov, V.A., Ural'ceva, N.N., 1968. *Linear and Quasilinear Equations of Parabolic Type.* American Mathematical Society. 7, 8, 10, 26, 29, 32

Lejay, A., 2000. *Méthodes probabilistes pour l'homogénéisation des opérateurs sous forme-divergence : cas linéaires et semi-linéaires.* PhD thesis, Université de Provence, Marseille, France. 20

Lejay, A., 2002a. BSDE driven by dirichlet process and semi-linear parabolic PDE. Application to homogenization. *Stochastic Process. Appl.* 97 (1), 1–39. 3, 9, 19, 24

Lejay, A., 2002b. Stochastic differential equations driven by processes generated by divergence form operators I: A Wong-Zakai theorem. INRIA/Institut Élie Cartan, Université Nancy I, preprint. 2, 13

Lunt, J., Lyons, T.J., Zhang, T.S., 1998. Integrability of functionals of Dirichlet processes, probabilistic representations of semigroups, and estimates of heat kernels. *J. Funct. Anal.* 153 (2), 320–342. DOI `10.1006/jfan.1997.3182`. 3

Lyons, T.J., Stoica, L., 1999. The limits of stochastic integrals of differential forms. *Ann. Probab.* 27 (1), 1–49. 2, 13

Ma, J., Yong, J., 1999. *Forward-backward stochastic differential equations and their applications*, vol. 1720 of Lecture Notes in Math. Springer-Verlag. 4

Orsina, L., Porzio, M.M., 1992. $L^\infty(Q)$-estimate and existence of solutions for some nonlinear parabolic equations. *Boll. Un. Mat. Ital. B* (7) 6 (3), 631–647. 33

Ōshima, Y., 1992a. On a construction of Markov processes associated with time dependent Dirichlet spaces. *Forum Math.* 4 (4), 395–415. 2

Ōshima, Y., 1992b. Some properties of Markov processes associated with time dependent Dirichlet forms. *Osaka J. Math.* 29 (1), 103–127. 2

Pardoux, É., 1999. BSDEs, weak convergence and homogenization of semilinear PDEs. In Clarke, F.H., Stern, R.J., editors, *Non Linear Analysis, Differential Equations and Control. Proc. Séminaire de Mathématiques Supérieures, Montréal 1998*, pp 503–549. Kluwer. 5, 23, 29

Porretta, A., 1999. Existence results for nonlinear parabolic equations via strong convergence of truncations. *Ann. Mat. Pura Appl.* (4) 177, 143–172. 33

Rozkosz, A., 1996a. Stochastic representation of diffusions corresponding to divergence form operators. *Stochastic Process. Appl.* 63 (1), 11–33. 2, 9, 13

Rozkosz, A., 1996b. Weak convergence of diffusions corresponding to divergence form operator. *Stochastics Stochastics Rep.* 57, 129–157. 9, 21

Rozkosz, A., 2002. Time-inhomogeneous diffusions corresponding to symmetric divergence form operators. *Probab. Math. Statist.* 22 (2), 231–252. 2

Rozkosz, A., 2003. Backward SDEs and Cauchy problem for semilinear equations in divergence form. *Probab. Theory Related Fields* 125 (3), 393–407. 3

Rozkosz, A., Słomiński, L., 1991. On existence and stability of weak solutions of multidimensional stochastic differential equations with measurable coefficients. *Stochastic Process. Appl.* 37 (2), 187–197. 5

Rozkosz, A., Slomiński, L., 1998. Extended convergence of Dirichlet processes. *Stochastics Stochastics Rep.* 65 (1–2), 79–109. 2

Stoica, L., 2003. A probabilistic interpretation of the divergence and BSDE's. *Stochastic Process. Appl.* 103 (1), 31–55. 3

Stroock, D.W., 1988. Diffusion semigroups corresponding to uniformly elliptic divergence form operator. In *Séminaire de Probabilités* XXII, vol. 1321 of Lecture Notes in Math., pp 316–347. Springer-Verlag. 8